DOCUMENT RESUME

ED 189 173 TM 800 355

AUTHOR	Cornish, Greg; Wines, Robin
TITLE	ACER Mathematics Profile Series: Number Test. (Test Booklet, Answer and Record Sheet, Score Key, and Teachers Handbook).
INSTITUTION	Australian Council for Educational Research, Hawthorn.
PUB DATE	80
NOTE	62p.
AVAILABLE FROM	Australian Council for Educational Research Limited, PO Box 210, Hawthorn, Victoria, Australia 3122 (6.30)
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Achievement Tests; Cognitive Development; Criterion Referenced Tests; *Developmental Stages; Foreign Countries; Item Analysis; *Latent Trait Theory; *Number Concepts; Quantitative Tests; Scaling; Scoring; Secondary Education; *Secondary School Mathematics; Testing; *Test Interpretation
IDENTIFIERS	*ACER Mathematics Profile Series; Australia; Piagetian Stages; *Rasch Model

ABSTRACT

The Number Test of the ACER Mathematics Profile Series, contains 30 items, for each of three suggested grade levels: 7-8, 8-9, and 9-10. Raw scores on all tests in the ACER Mathematics Profile Series (Number, Operations, Space and Measurement) are converted to a common scale called MAPS, a major feature of the Series. Based on the Rasch Model, MAPS makes it possible to estimate the likelihood of a student with a known ability being able to correctly answer an item of known difficulty. Broadly speaking, MAPS also makes it possible to measure the latent variable underlying the Series: mathematics development. Expected mastery levels for test items are related to four Piagetian developmental stages: early concrete operational; late concrete operational; early formal operational; and late formal operational. The objectives which are suggested are process, rather than content, oriented. Based loosely upon Bloom's Taxonomy, they relate mathematics development to such cognitive processes as knowledge, operations, classification, ordination, relevance, creativity, and problem solving. (CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED189173

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

J. P. Keeves

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Contents

	Page
Unit I	1
Unit II	5
Unit III	11

ACER MATHEMATICS PROFILE SERIES

NUMBER TEST

Unit I

Directions

This unit contains 30 questions. Although some questions may not be directly related to the work you have done in class, you should try to work out the correct answer.

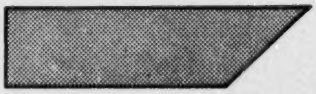
In each question the alternative answers are represented by the letters **A, B, C, D,** or **E.** Mark your chosen answer by circling the appropriate letter on the Answer and Record Sheet.

The following practice questions will show you how to answer the questions on the test.

Wait until you are told how to answer the questions before going on.

Practice Questions

P1



The measure of the shaded region in this diagram is called the

A area. **C** circumference.
B perimeter. **D** volume.

Since area is the correct answer, the corresponding letter **A** has been marked on the answer sheet.

P2 A person's average walking speed is 6 kilometres per hour. Which one of the following is the distance the person would travel in 1 hour?

A 2 km **C** 6 km
B 3 km **D** 12 km
E 24 km

Please do not make any marks in this booklet

BACKGROUND INFORMATION FOR TEACHERS

Test materials in the *ACER Mathematics Profile Series* are designed so that teachers may monitor students' mathematical development throughout primary and secondary schooling. This is achieved by converting raw scores on any of the tests to a common scale called the MAPS scale. Conversion tables for this test are incorporated in the *ACER Mathematics Profile Series – Number Test Teachers Handbook.* The Handbook discusses interpretive procedures concerning the likely mastery of all items on the MAPS scale. These interpretations enable teachers to identify a range of suitable learning experiences, in relation to each student.

To assist in the calculation and use of mastery levels the *Answer and Record Sheet* has a specially designed student record section and the score key displays a *mastery profile 'cursor'.*

The suggested testing time for a *Number Test Unit* is about 30 minutes.

Published by
The Australian Council for Educational Research Limited,
Radford House, Frederick Street, Hawthorn, Victoria, 3122

Copyright © ACER 1980

1 The temperature at 10 a.m. was 12.6 °C. At 11.30 a.m. it was 18.9 °C. Which one of the following gives the increase in temperature?

A 6.3 C° **C** 30.5 C°
B 6.7 C° **D** 31.5 C°

2 An overseas cruise leaves Australia in the middle of the 4th month of every year. The cruise lasts 10 months. In what month of the following year does it end?

A 2nd **C** 8th
B 4th **D** 10th

3 Which one of the following gives 659 rounded to the nearest hundred?

A 600 **C** 660
B 650 **D** 700

4 If m stands for any number, 4 × m can be written as

A m + m + m + 4.
B m ÷ 4.
C m + m + m + m.
D m + 4.

5 Four lines are related as follows:

Line I is twice as long as line II

Line II is $\frac{3}{4}$ the length of line III

Line III is half as long as line IV.

Which line is the longest?

A I **C** III
B II **D** IV

6 What is the average of the set of numbers 5, 7, 12?

A 5 **C** 8
B 7 **D** 9

7 Jimmy has eaten one piece of pie, as shown in the diagram.

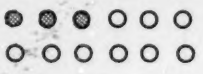
He decides to cut the rest of the pie into pieces the same size as the piece he ate. How many pieces of pie will there be now?

A 4 **C** 6
B 5 **D** 7
E 9

8 Which one of the following diagrams does **not** have $\frac{1}{3}$ shaded?

A

B



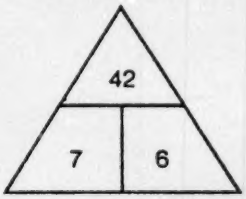
C

D

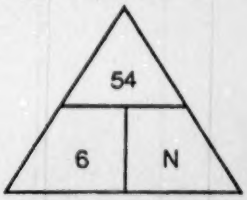
9 A boy wrote 166 × 53 in mistake for 168 × 53. By how much was his answer too small?

A 2 **C** 332
B 106 **D** 336

10 The three numbers in Triangle I are related.



Triangle I



Triangle II

If the pattern in Triangle I is repeated in Triangle II, which one of the following would give the value of N?

A 6 + N = 54 **C** N × 54 = 6
B 54 + 6 = N **D** 54 − N = 6

11 In a magic square, the entries in any row, column, or diagonal always add up to the same number. Some entries in this 3 × 3 magic square are missing.

5	19	x
	11	
		17

Which one of the following statements is true about x?

A x must stand for 7
B x must stand for 9
C x must stand for 24
D x must stand for 33
E There is insufficient information to find the value of x.

The Australian Council for Educational Research Limited, Hawthorn, Victoria, 3122. Copyright © ACER 1980 ACER Mathematics Profile Series – Number Test – Unit I

12 A team has won 3 of the 6 games already played. If the team wins the next four games, what percentage of the games played will it then have won?

A	40%	**C**	70%
B	50%	**D**	75%

13 This number track is followed from left to right. Each circle represents the number resulting from the operation shown on the previous section of track.

The number that replaces Y would be

A 48.
B 36.
C 23.
D 19.

14 A jellybean machine works at a constant rate making 24 000 jellybeans every 24 hours. How many jellybeans would the machine make in 4 hours?

A	1000	**C**	4000
B	2000	**D**	6000

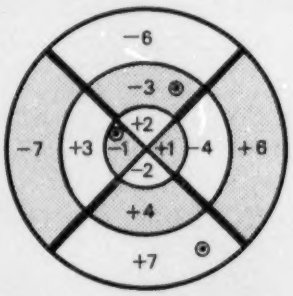
15 Tom has 9 marbles. He plays two games of marbles. He loses three marbles in the first game, and another two in the second.
Which one of the following gives the number of marbles he had left?

A	(9 − 3) + 2	**C**	(9 − 3) − 2
B	9 − (3 − 2)	**D**	9 + (3 − 2)

16 If half an apple is shared equally be three people, what fraction of a whole apple does each person receive?

A	$\frac{1}{3}$	**C**	$\frac{1}{6}$
B	$\frac{1}{4}$	**D**	$\frac{1}{8}$

17 The diagram represents an 'integer' dartboard. A score is given by the region in which a dart lands.



If a person threw 3 darts which landed on the positions shown by the ◉'s, what was the sum of his scores?

A	−3	**C**	−11
B	+3	**D**	+11

18 A can of pineapple weighs 454 gram. What is the total weight in kilogram, given 1 gram = 0.001 kilogram, of 10 cans of pineapple?

A	0.0454 kg	**C**	4.54 kg
B	0.454 kg	**D**	45.4kg
		E	454 000 kg

19 If □ stands for any odd number, then $(2 \times \square) - 1$

A must be odd.
B must be even.
C can be odd or even.

20 For every 3 bottles collected, a girl receives 7 cents. If she collected 210 bottles, which expression would give the total amount of money she received?

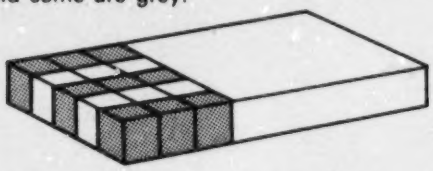
A	$\frac{210 \times 3}{7}$ cents	**C**	$\frac{210 \times 7}{3}$ cents
B	$\frac{3 \times 7}{210}$ cents	**D**	$\frac{210}{3 \times 7}$ cents

21 The manufacturing cost of a radio is $120. If 25 per cent sales tax is then added to this cost, what is the total cost of the radio?

A	$90	**C**	$145
B	$95	**D**	$150

The Australian Council for Educational Research Limited, Hawthorn, Victoria, 3122. Copyright © ACER 1980 ACER Mathematics Profile Series – Number Test – Unit I

22 The diagram shows a partly covered box containing 40 cubes altogether. Some of the cubes are white and some are grey.



If the pattern of the cubes continues under the cover, how many of the 40 cubes are grey?

A 9 **C** 24
B 20 **D** 30

23 A pole 100 centimetre in length is painted as follows; 0.25 of its length is painted white, 0.25 of its length is painted red, 0.4 of its length is painted yellow, and the remainder is painted blue.

The length of pole painted blue is

A 10 cm. **C** 25 cm.
B 20 cm. **D** 40 cm.

24 What value does the digit '9' have, after 27.94 is multiplied by 10?

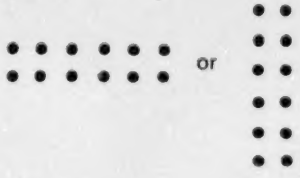
A 9 tens **C** 9 tenths
B 9 units **D** 9 hundredths

25 A tank, one-quarter full of water, requires an extra 2500 litres to **half** fill it. How many litres of water does the tank hold when full?

A 5000 **C** 10 000
B 7500 **D** 25 000
E 100 000

26 'Rectangular' patterns can be formed from different numbers of dots if each pattern is rectangular in shape and one side has at least 2 dots.

For example, 12 dots can only form two different patterns as shown.

pattern 1  or

pattern 2 or

Which one of the following number of dots will form only **one** rectangular pattern?

A 16 **C** 20
B 18 **D** 27
E 28

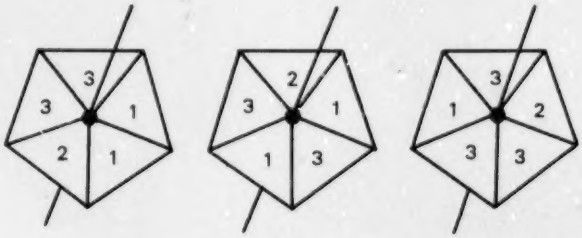
27 A prize of $21 000 was divided between three people. If Peter received twice as much as John, and Jenny received twice as much as Peter, Peter's share was

A $2000. **C** $6000.
B $3000. **D** $12 000.

28 When $(-7 - 2) + 3$ is divided by a number the answer is $+2$.
What is the number?

A -3 **C** $+\frac{1}{3}$
B $-\frac{1}{3}$ **D** $+3$

29 Three different spinners, I, II, III, are numbered as shown. Each spinner is spun once, and the score is the number on the edge which comes to rest on the table.



I II III

If a person needed to spin a '3' to win a game, which spinner would give the greatest chance of winning?

A spinner I
B spinner II
C spinner III
D any spinner

30 If $\frac{5}{7}n = 1$, then the value of n is

A $\frac{7}{5}$. **C** $\frac{7}{1}$.
B $\frac{7}{7}$. **D** $\frac{5}{7}$.

The Australian Council for Educational Research Limited, Hawthorn, Victoria, 3122. Copyright © ACER 1980 ACER Mathematics Profile Series – Number Test – Unit 1

ACER MATHEMATICS PROFILE SERIES

NUMBER TEST

Unit II

Directions

This unit contains 30 questions. Although some questions may not be directly related to the work you have done in class, you should try to work out the correct answer.

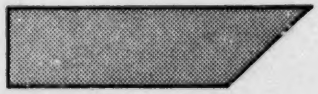
In each question the alternative answers are represented by the letters **A, B, C, D,** or **E.** Mark your chosen answer by circling the appropriate letter on the Answer and Record Sheet.

The following practice questions will show you how to answer the questions on the test.

Wait until you are told how to answer the questions before going on.

Practice Questions

P1



The measure of the shaded region in this diagram is called the

A area.
B perimeter.
C circumference.
D volume.

Since area is the correct answer, the corresponding letter **A** has been marked on the answer sheet.

P2 A person's average walking speed is 6 kilometres per hour. Which one of the following is the distance the person would travel in 1 hour?

A 2 km
B 3 km
C 6 km
D 12 km
E 24 km

Please do not make any marks in this booklet

BACKGROUND INFORMATION FOR TEACHERS

Test materials in the *ACER Mathematics Profile Series* are designed so that teachers may monitor students' mathematical development throughout primary and secondary schooling. This is achieved by converting raw scores on any of the tests to a common scale called the MAPS scale. Conversion tables for this test are incorporated in the *ACER Mathematics Profile Series – Number Test Teachers Handbook.* The Handbook discusses interpretive procedures concerning the likely mastery of all items on the MAFS scale. These interpretations enable teachers to identify a range of suitable learning experiences, in relation to each student.

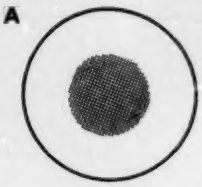
To assist in the calculation and use of mastery levels the *Answer and Record Sheet* has a specially designed student record section and the score key displays a *mastery profile 'cursor'.*

The suggested testing time for a *Number Test Unit* is about 30 minutes.

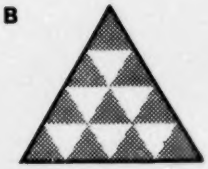
Published by
The Australian Council for Educational Research Limited,
Radford House, Frederick Street, Hawthorn, Victoria, 3122

Copyright © ACER 1980

1 Which one of the following shapes has half of its area shaded?

A 

C

B 

D

2 In the equation △ stands for an operational sign.

$$5 \triangle 4 = 4 \triangle 5$$

Which of the following operational signs can be used to make the equation true?

A + only
B × only
C + or ×
D × or +
E + or −

3 A school has one teacher for every 22 pupils. If the school has 88 teachers, which one of the following is the best estimate of the number of pupils attending the school?

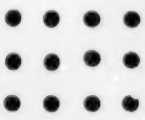
A 352
B 700
C 1500
D 1900

4 The temperature at 8 a.m. was 7.8 °C and by 2 p.m. it had risen by 16.6 C degrees.
Which one of the following was the temperature at 2 p.m.?

A 8.8 °C
B 16.6 °C
C 23.14 °C
D 24.4 °C

5 'Rectangular' patterns can be formed from different numbers of dots so that each pattern is at least 2 dots wide.
For example, 12 dots form these two patterns.

pattern 1


pattern 2

Which one of the following numbers of dots will **not** make a rectangular pattern?

A 15
B 21
C 29
D 39

6 How many pieces of tape 12 centimetre (0.12 metre) long can be cut from a 6 metre tape?

A 2
B 18
C 50
D 200

7 The operation table for {6, 7, 8, 9} under the operation ⊛ is shown in the table.

⊛	6	7	8	9
6	8	9	6	7
7	9	6	7	8
8	6	7	8	9
9	7	8	9	6

Which one of the following is the value of 7 ⊛ 9?

A 6
B 7
C 8
D 9

8 A girl sent one card each to 21 friends. Each card cost 25 cents, and the postage for each card was 18 cents.
Which one of the following gives the total cost?

A (21 + 25) + 18 cents
B 21 × (25 + 18) cents
C 21 + (25 × 18) cents
D (21 + 25) × (21 + 18) cents

9 Consider the operation ▽ such that
$p \triangledown q = 2 \times p \times q$.
For example, $3 \triangledown 4 = 2 \times 3 \times 4$
$= 24$

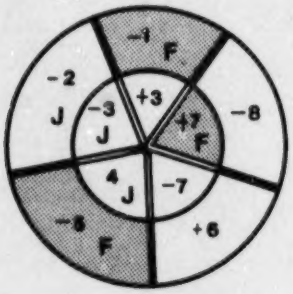
Which one of the following is equal to $6 \triangledown 3$?

A 9
B 15
C 18
D 36

The Australian Council for Educational Research Limited, Hawthorn, Victoria, 3122. Copyright © ACER 1980 ACER Mathematics Profile Series – Number Test – Unit II

10 The diagram represents an 'integer' dartboard. Two players each threw three darts.

Fred's darts landed in the regions marked by F, and Jodie's darts landed in the regions marked by J.



A player's scores are summed and the player with the highest positive sum is the winner. Which one of the following statements is true?

A Fred won by 1 point.
B Fred won by 2 points.
C Jodie won by 1 point.
D Jodie won by 2 points.
E It was a draw, since both players threw the same score.

11 Each time a person writes a cheque, the amount of the cheque is subtracted from their account. If a person writes a cheque for more than the amount they have in credit, they are overdrawn.
Mary's account is \$630.43 in credit. She writes a cheque for \$207. Which one of the following gives the state of her account?

A \$837.43 in credit
B \$837.43 overdrawn
C \$423.43 overdrawn
D \$423.43 in credit

12 After selling $\frac{3}{4}$ of his papers, a newsboy had 60 left.

How many papers did he sell?

A 60 **C** 180
B 80 **D** 240

13 The average of the set of four numbers 4, 5, p, q is 5. Which one of the following pairs of numbers could **not** be p and q?

A 2 and 9 **C** 4 and 8
B 3 and 8 **D** 5 and 6
E 7 and 4

14 Which one of the following is **not** equal to 2^8?

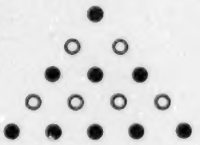
A $2^2 \times 2^6$ **C** $2^2 \times 2^2 \times 2^2 \times 2^2$
B $(2^2)^6$ **D** $\frac{2^{10}}{2^2}$

15 The sum of the first four odd numbers is equal to the square of 4, or 16.
That is,

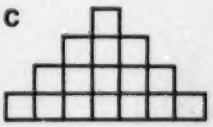
$$1 + 3 + 5 + 7 = 16.$$

Which one of the following does **not** illustrate this relationship?

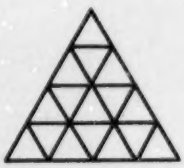
A



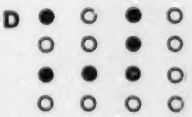
C



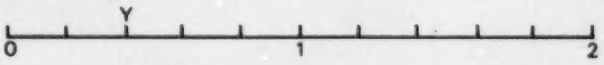
B



D



16 The diagram shows part of a number line.



Which one of the following represents the point Y?

A 1.2 **C** 0.2
B 0.4 **D** 0.04

17 The table gives some information on smokers and lung cancer.

		Smokers		
		heavy	light	total
Cause of Death	lung cancer	19	1	20
	other	181	799	980
	total	200	800	1000

Of all those who died of lung cancer, what fraction were heavy smokers?

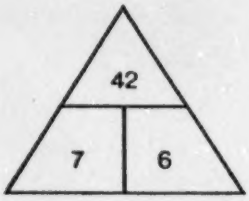
A $\frac{19}{20}$ **C** $\frac{19}{200}$
B $\frac{19}{181}$ **D** $\frac{19}{1000}$
E $\frac{20}{1000}$

The Australian Council for Educational Research Limited, Hawthorn, Victoria, 3122. Copyright © ACER 1980 ACER Mathematics Profile Series – Number Test – Unit II

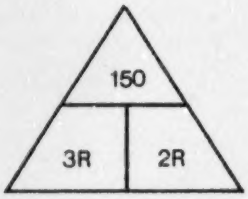
18 A refrigerator has a marked price of $250, but is sold with a discount for $200. The percentage discount offered on the marked price was

A 5%. **C** 25%.
B 20%. **D** 50%.

19 The three numbers in Triangle I are related.



Triangle I



Triangle II

If the pattern in Triangle I is repeated in Triangle II, what is the value of R?

A 5 **C** 25
B 10 **D** 30

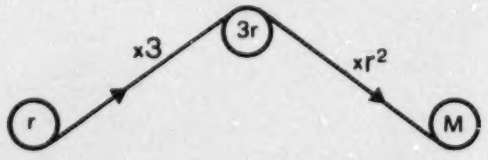
20 If $n^{\triangle} \times n^{*} = n^{\square}$, then $\square$ equals

A $\triangle + *$. **C** $\triangle \div *$.
B $\triangle \times *$. **D** $\triangle - *$.

21 Stuart bought some books. If he had paid for them with $2 notes, he would have paid out 6 more notes than if he had paid with $5 notes. How much did the books cost?

A $15 **C** $25
B $20 **D** $30

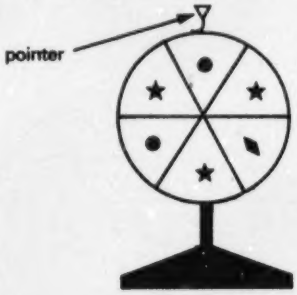
22 This number track is followed from left to right. Each circle represents the number resulting from the operation shown on the previous section of track.



Written in simplest form, which one of the following is M?

A $3r^2$ **C** $6r$
B $3r^3$ **D** $6r^2$
E $3r + r^2$

23 The diagram shows a spinning wheel. This spinning wheel has been made so that the pointer always finishes up pointing to a symbol.



If the wheel has been spun and then stops, which symbol is the pointer most likely to point towards?

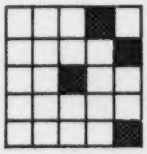
A ★

B ♦

C ●

D ★, ♦, ●, as they all have equal chance

24 The diagram shows a square divided into equal regions.



What **percentage** of the square is shaded?

A 4% **C** 16%
B 8% **D** 25%

25 If m is a natural number, the solution set of $\frac{m}{5} < \frac{5}{6}$ is

A {1}. **C** {1, 2, 3}.
B {1, 2}. **D** {1, 2, 3, 4}.

26 Four cans of soup were tested for water content. Each can was the same size and it was found that

can I contained 30% water.
can II contained $\frac{1}{3}$ water.
can III contained $\frac{1}{4}$ water.
can IV contained 0.31 water.

Which can of soup contained the most water?

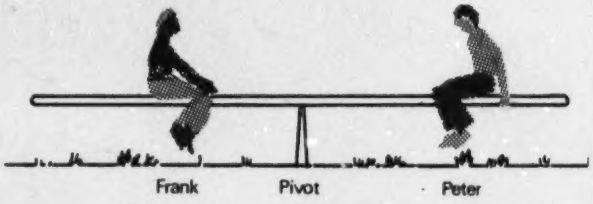
A I **C** III
B II **D** IV

The Australian Council for Educational Research Limited, Hawthorn, Victoria, 3122. Copyright © ACER 1980 ACER Mathematics Profile Series – Number Test – Unit II

27 In a class of 20 students, 35 per cent are boys and the remainder are girls. How many girls are there in the class?

A 6
B 7
C 13
D 14
E 15

28 Peter and Frank sit on a see-saw as shown in the diagram. Peter is three-quarters the distance from the pivot to his end of the see-saw, and Frank is half the distance from the pivot to his end of the see-saw.



If the see-saw is horizontally balanced, which one of the following statements is true?

A Peter is $\frac{2}{3}$ Frank's weight.

B Peter is $\frac{3}{8}$ Frank's weight.

C Frank is $2\frac{2}{3}$ Peter's weight.

D Peter is $1\frac{1}{2}$ times Frank's weight.

E Frank and Peter are the same weight.

29 The symbols $\square$, $\triangle$, $\ominus$, stand for natural numbers.
Which one of the following statements is always true?

A $\triangle + (\square \times \ominus) = (\triangle + \square) \times (\triangle + \ominus)$

B $\triangle \times (\square \times \ominus) = (\triangle \times \square) \times (\triangle \times \ominus)$

C $\triangle + (\square \times \ominus) = (\triangle + \square) + (\triangle + \ominus)$

D $\triangle \times (\square + \ominus) = (\triangle \times \square) + (\triangle \times \ominus)$

30 If n is greater than 3, which one of the following is the smallest?

A $\frac{3}{n}$
B $\frac{3}{n+1}$
C $\frac{n}{3}$
D $\frac{n+1}{3}$

The Australian Council for Educational Research Limited, Hawthorn, Victoria, 3122. Copyright © ACER 1980 ACER Mathematics Profile Series – Number Test – Unit II

ACER MATHEMATICS PROFILE SERIES

NUMBER TEST

Unit III

Directions

This unit contains 30 questions. Although some questions may not be directly related to the work you have done in class, you should try to work out the correct answer.

In each question the alternative answers are represented by the letters **A, B, C, D,** or **E.** Mark your chosen answer by circling the appropriate letter on the Answer and Record Sheet.

The following practice questions will show you how to answer the questions on the test.

Wait until you are told how to answer the questions before going on.

Practice Questions

P1

The measure of the shaded region in this diagram is called the

A area.
B perimeter.
C circumference.
D volume.

Since area is the correct answer, the corresponding letter **A** has been marked on the answer sheet.

P2 A person's average walking speed is 6 kilometres per hour. Which one of the following is the distance the person would travel in 1 hour?

A 2 km
B 3 km
C 6 km
D 12 km
E 24 km

Please do not make any marks in this booklet

BACKGROUND INFORMATION FOR TEACHERS

Test materials in the *ACER Mathematics Profile Series* are designed so that teachers may monitor students' mathematical development throughout primary and secondary schooling. This is achieved by converting raw scores on any of the tests to a common scale called the MAPS scale. Conversion tables for this test are incorporated in the *ACER Mathematics Profile Series – Number Test Teachers Handbook*. The Handbook discusses interpretive procedures concerning the likely mastery of all items on the MAPS scale. These interpretations enable teachers to identify a range of suitable learning experiences, in relation to each student.

To assist in the calculation and use of mastery levels the *Answer and Record Sheet* has a specially designed student record section and the score key displays a *mastery profile 'cursor'*.

The suggested testing time for a *Number Test Unit* is about 30 minutes.

Published by
The Australian Council for Educational Research Limited,
Radford House, Frederick Street, Hawthorn, Victoria, 3122

Copyright © ACER 1980

1 The diagram shows part of a number line.

Which one of the following represents a point in the shaded part of this number line?

A $\frac{1}{2}$ **C** $\frac{4}{5}$

B $\frac{3}{4}$ **D** $\frac{1}{8}$

2 Bryce cut 7.5 metre of braid into pieces each 0.75 metre in length. How many 0.75 metre lengths of braid did he have?

A 8 **C** 10

B 9 **D** 11

3 If $p = q \times r$, and q and r represent negative integers, then

A p is negative for all q and r.

B p is positive for all q and r.

C p is positive only if q is greater than r.

D p is negative only if q is less than r.

4 Consider the operation $\multimap$ such that

$p \multimap q = 2p + q$

For example, $5 \multimap 6 = (2 \times 5) + 6$
$= 16$

Which one of the following is equal to $4 \multimap (-3)$?

A 2 **C** −11

B 5 **D** −24

5 A carton holds 60 tennis balls, and a canister holds 3 tennis balls. How many canisters would be filled from 2 full cartons of tennis balls?

A 40 **C** 120

B 60 **D** 180

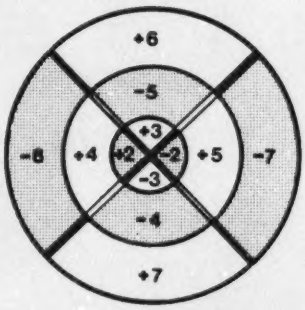
6 The average of the set of three numbers, 2, 3, n, is 4.
Which one of the following numbers is n?

A 4 **C** 8

B 7 **D** 9

E 12

7 The diagram represents an 'integer' dartboard. A score is given by the region in which a dart lands.



In an 'integer' dart game two players scored as follows;

Player I +6, +3, −5
Player II +6, +5, N

The player with the nignest positive sum of scores is the winner. If each player's scores are totalled, Player II wins the game by one point.
Which one of the following is the value of N?

A −7 **C** + 6

B −6 **D** + 7

8 In Robin's family there were 2 parents, 4 grandparents, and 8 great-grandparents. How many great-great-grandparents did she have?

A 2^3 **C** 12

B 2^4 **D** 14

9 Which one of the following is the largest element of the set $\{222, 22^2, (2^2)^2, 2^{22}\}$?

A 222 **C** $(2^2)^2$

B 22^2 **D** 2^{22}

10 If m and p are natural numbers and m is less than p, which one of the following is always true?

A $\frac{m}{p}$ is less than 1.

B $\frac{m}{p}$ is greater than 1.

C $\frac{m}{p}$ can be equal to 1.

D $\frac{m}{p}$ can be both less than or greater than 1.

The Australian Council for Educational Research Limited, Hawthorn, Victoria, 3122. Copyright © ACER 1980 ACER Mathematics Profile Series – Number Test – Unit III

11 The symbols □, △, ⊖ stand for natural numbers.

If $\frac{\square}{\triangle} = \frac{\square \times \ominus}{M}$, then M is equal to

A □ × ⊖.
B □ × △.
C ⊖ × □.
D ⊖ × △.

12 Steven's account is overdrawn by $314.86. He withdraws $85.33 from his account, and pays $280 into his account.
Which one of the following is a correct statement of his account at present?

A $120.19 overdrawn
B $400.19 overdrawn
C $50.47 in credit
D $680.19 in credit

13 A man, waiting to use a telephone, walks up and down a footpath which runs east-west. Starting from the telephone box, he walked 5 metres east, 10 metres west, 15 metres east, and 20 metres west before the telephone box was vacant. How far must he now walk to reach the telephone box?

A 5 metres west
B 10 metres east
C 10 metres west
D 20 metres east
E neither **A**, nor **B**, nor **C**, nor **D**

14 A waiter wants to make up 35 litres of drink using 2 parts of cordial syrup to every 5 parts of water. How many litres of cordial syrup are needed?

A 5
B 10
C 14
D 16

15 When the number 5 is multiplied by all the natural numbers less than it, e.g. 5 × 4 × 3 × 2 × 1, the mathematical shorthand used to represent this product is 5!, called 'factorial 5'.

Similarly 10! = 10 × 9 × 8 × 7 × 6 × 5 × 4 × 3 × 2 × 1.

If the value of 5! is 120, then the value of 7! can be given by

A 7 × 120.
B 42 × 120.
C 76 × 120.
D 13 × 120.

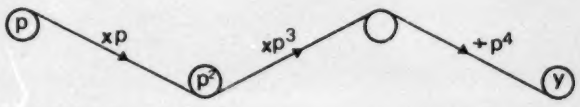
16 Suppose that the Babylonians, who used 60 as their base number, had developed a system similar to our 'percentages' in which they expressed fractions with denominator sixty as 'persixtages'.

For example, $\frac{23}{60}$ would be called '23 persixt'.

In our system, Babylonian '6 persixt' would be equivalent to

A 1%.
B 6%.
C 10%.
D 60%.

17 This number track is followed from left to right. Each circle represents the number resulting from the operation shown on the previous section of track.



Written in the simplest form, which one of the following represents Y?

A $\frac{5}{4}p$
B p
C $p^{\frac{5}{4}}$
D $p^{\frac{3}{2}}$
E $\frac{3}{2}p$

18 There are 21 boxes, and one ball is placed into each of an odd number of the boxes. Into half of the rest of the boxes are placed two balls per box.
Which one of the following is the total number of balls in the boxes?

A 15
B 21
C 31
D The total number varies according to the odd number of boxes chosen.

19 A mathamoeba reproduces by trinary fission each day. That is, a mathamoeba splits into 3 each day. Suppose on Day 0 there was one mathamoeba, then on Day 1 there would be 3 mathamoeba. How many mathamoeba would there be on Day 5?

A 3 × 4
B 3 × 5
C 3^4
D 3^5
E 3^6

The Australian Council for Educational Research Limited, Hawthorn, Victoria, 3122. Copyright © ACER 1980 ACER Mathematics Profile Series – Number Test – Unit III

20 In the expression $R = a \times b^2$, a and b represent negative numbers. Which one of the following statements about R is true?

A R is always negative.

B R is always positive.

C R is positive if b is greater than a.

D R is negative if b is less than a.

21 The relation $n^p = m$ can be written as $\log_n m = p$. If $\log_n 32 = p$, then

A $p = 5, n = 2$.

B $p = 2, n = 5$.

C $p = 4, n = 8$.

D $p = 8, n = 16$.

E $p = 2, n = 16$.

22 In 'clock' 4 arithmetic only the numbers 1, 2, 3, 4 are used. The operation table for 'clock' 4 multiplication on the set {1, 2, 3, 4} is shown in the table.

x	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

Which one of the following is equal to $(4 \times 3) \times 3$?

A 1

B 3

C 4

D 12

23 If one of two factors of 3^{-4} is 3^3, which one of the following is the other factor?

A 3^{-7}

B 3^{-1}

C 3

D 3^7

24 A certain material shrinks 0.5 per cent when wet. Expressed as a proper fraction this is

A $\frac{1}{2}$

B $\frac{1}{20}$

C $\frac{1}{200}$

D $\frac{1}{2000}$

25 A golf ball manufacturer guarantees that not more than $3\frac{1}{2}$ per cent of the golf balls he produces are defective. From a sample of 14 000 golf balls, how many would he expect to be **not** defective?

A 420

B 490

C 13 020

D 13 510

E 13 790

26 Which one of the following is the best approximation for $\frac{2}{3}$?

A 0.667

B 0.67

C 0.66

D 0.7

27 In a competition five of the cash prizes, each of $8, were awarded to five girls. The remaining 20 per cent of the prize money was shared equally by four boys.
Which one of the following gives the amount each boy received?

A $10.00

B $8.00

C $2.50

D $2.00

28 When a whole number k is divided by 9 the remainder is 5. What would be the remainder if 2k were divided by 9?

A 1

B 5

C 10

D neither **A**, nor **B**, nor **C**.

29 A full can of kerosene weighs 8 kg. If half of the kerosene in the can was poured out the weight would be 4.5 kg. What is therefore the weight of the empty can?

A 0.25 kg

B 0.5 kg

C 1.0 kg

D 1.75 kg

E There is insufficient information to decide.

30 $\frac{a}{b} + \frac{c}{d}$ is equivalent to

A $\frac{a + c}{b + d}$

B $\frac{a + c}{bd}$

C $\frac{ad + bc}{bd}$

D $\frac{ac}{bd}$

The Australian Council for Educational Research Limited, Hawthorn, Victoria, 3122. Copyright © ACER 1980 ACER Mathematics Profile Series – Number Test – Unit III

ACER MATHEMATICS PROFILE SERIES

NUMBER TEST

Answer and Record Sheet

Unit I II III

Name .....

Class .....

Date of Testing ...../...../.....

Date of Birth ...../...../.....

Age .....year .....months

NOTE: This *Answer and Record Sheet* may be used for any one of the three Units. Many questions have only four alternative answers.

MASTERY LEVELS

90%		−10
75%		−5
50%		*MAPS ability*
25%		+5
10%		+10

Raw Score .....

Scaled Error .....

Practice Questions	P1	A	B	C	D	E	P2	A	B	C	D	E
1		A	B	C	D	E	16	A	B	C	D	E
2		A	B	C	D	E	17	A	B	C	D	E
3		A	B	C	D	E	18	A	B	C	D	E
4		A	B	C	D	E	19	A	B	C	D	E
5		A	B	C	D	E	20	A	B	C	D	E
6		A	B	C	D	E	21	A	B	C	D	E
7		A	B	C	D	E	22	A	B	C	D	E
8		A	B	C	D	E	23	A	B	C	D	E
9		A	B	C	D	E	24	A	B	C	D	E
10		A	B	C	D	E	25	A	B	C	D	E
11		A	B	C	D	E	26	A	B	C	D	E
12		A	B	C	D	E	27	A	B	C	D	E
13		A	B	C	D	E	28	A	B	C	D	E
14		A	B	C	D	E	29	A	B	C	D	E
15		A	B	C	D	E	30	A	B	C	D	E

TM 800355 (a-f) (1)

The Australian Council for Educational Research Limited, Radford House, Hawthorn, Victoria 3122

JB·ABCD/8765·4321

ACER MATHEMATICS PROFILE SERIES

NUMBER TEST

Score Key (on both sides)

mastery profile 'cursor' on other side

The other components of the *Number Test* are
Number Test Test Booklet
Number Test Answer and Record Sheet
Number Test Teachers Handbook

The Australian Council for Educational Research Limited
Radford House, Frederick St., Hawthorn. 3122

Copyright © ACER 1980

18

UNIT I

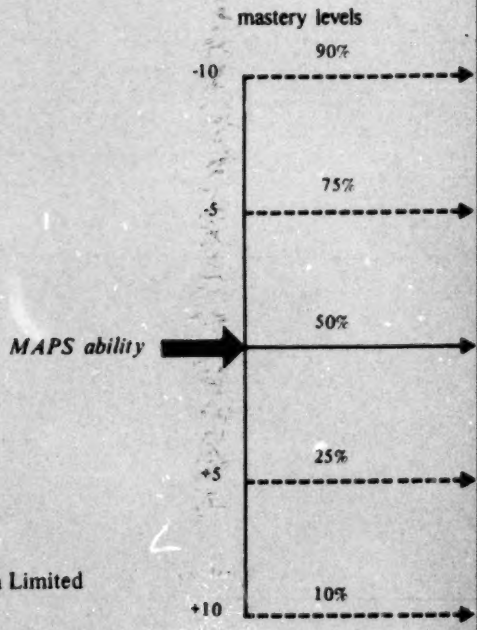
1	A	30	A
2	A	29	C
3	D	28	A
4	C	27	C
5	D	26	D
6	C	25	C
7	D	24	B
8	B	23	A
9	B	22	C
10	B	21	D
11	B	20	C
12	C	19	A
13	D	18	C
14	C	17	B
15	C	16	C

UNIT II

1	C	15	A
2	D	14	B
3	D	13	C
4	D	12	C
5	C	11	D
6	C	10	B
7	C	9	D
8	B	8	B
9	D	7	C
10	B	6	B
11	D	5	A
12	C	4	B
13	C	3	A
14	B	2	B
15	A	1	B

UNIT III

1	A	2	C	3	B	4	B	5	A	6	B	7	B	8	B	9	D	10	A	11	D	12	A	13	B	14	B	15	B	16	C
30	C	29	C	28	A	27	C	26	A	25	D	24	C	23	A	22	A	21	A	20	A	19	D	18	B	17	B				



MASTERY PROFILE 'CURSOR'

The mastery profile 'cursor' is used in conjunction with the *Number item profiles* given in the *Number Test* Teachers Handbook.

Place the arrow ➡ against a student's *MAPS ability*. Items which lie in suitable mastery ranges can be identified.

Interpretation procedures are discussed in the *Number Test* Teachers Handbook.

The Australian Council for Educational Research Limited
Radford House, Frederick St., Hawthorn. 3122

Copyright © ACER 1980

JB-ABCD/8765-4321

ACER MATHEMATICS PROFILE SERIES

NUMBER TEST

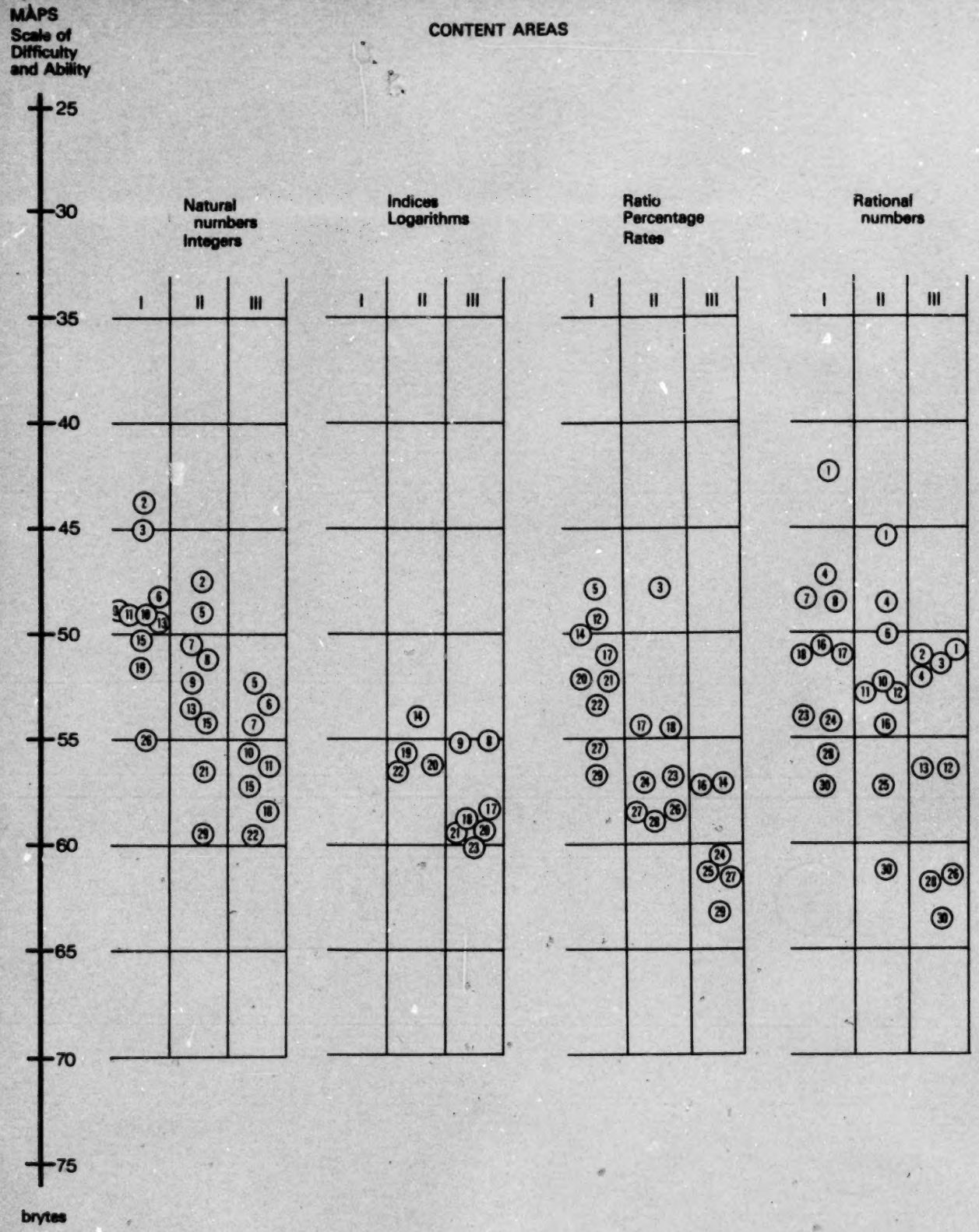
Teachers Handbook

- Administration and scoring
- Use and interpretation
- The Rasch measurement approach
- Broader interpretations

TM 800 355 (4 of 4)

Number Test — item content profile overleaf.

This profile can be used in conjunction with the conversion tables given in Section IV.



Number Test – item profile: Content Areas

Experimental edition

ACER MATHEMATICS PROFILE SERIES

NUMBER TEST

Teachers Handbook

Greg Cornish and Robin Wines

Administration and scoring
Use and interpretation
The Rasch measurement approach
Broader interpretations

The Australian Council for Educational Research Limited,
Hawthorn, 3122, Victoria.

Published by: The Australian Council for Educational Research Limited
Radford House, Frederick Street, Hawthorn, Victoria 3122.

Typeset 10/12 Times Roman by: Savage and Co. Pty. Ltd.,
36 Cordelia Street, Fortitude Valley, 4006, Queensland.

Printed and bound by: Jenkin Buxton Printers Pty. Ltd.,
1 Abbotsford Street, West Melbourne, Victoria 3003.

Copyright © ACER 1980

This handbook in the *ACER Mathematics Profile Series* is released as an experimental edition, and comment is invited on the material presented.

ACKNOWLEDGMENT

ACER thanks the teachers and students from the following schools for their participation in the trial testing program

Brackenridge Primary School
Debney Park High School
Ibrox Park Boys High School
Mt. Gravatt State High School
North Sydney Girls High School
Oak Park High School
Sandgate District High School
Upper Mt. Gravatt Primary School
Williamstown Technical High School
Yarra Valley Church of England School

No part of this publication may be reproduced in any form without permission from the publisher.

The *Number Test* comprises the *Number Test* test booklet, answer and record sheet, score key, and teachers handbook.
The *Number Test* is part of the *ACER Mathematics Profile Series*.

Foreword

The publication of this handbook and the accompanying test marks a potentially important development in Australian Mathematics Education. The *ACER Mathematics Profile Series* contains a number of innovative features aimed at advancing mathematics teaching and learning. The most significant of these is the application of the Rasch approach to develop the MAPS scale and to provide teachers with a flexible testing system which relates item difficulty and student ability.

This project began in 1974 and was supported in its early stages by grants in two years of $11 000 per annum from the Victorian Education Department and more recently by the secondment of Mr. Greg Cornish from that Department for a period of three years.

The items in the *Operations Test* are based on those developed by Professor K. F. Collis and ACER gratefully acknowledges this contribution to the program.

In 1977 a Committee was established to advise ACER staff on the publication of the *Profile Series.* This Committee consisted of:

Mr. M. A. Clements (Monash University)
Mr. J. F. Izard (SCV Coburg)
Mr. W. Newton (Curriculum & Research Branch, Victoria)
Ms. R. Wines (Yarra Valley School)

In 1978 the Committee consisted of:

Mr. M. A. Clements (Monash University)
Mr. W. Newton (Curriculum & Research Branch, Victoria)
Ms. R. Wines (Yarra Valley School)

The *ACER Mathematics Profile Series* is a tribute to the two authors. Ms. Robin Wines deserves the credit for the development of the program and its organization over the initial years. Ms. Wines at first worked alone but was joined in January 1976 by Mr. Cornish. Mr. Cornish has prepared this handbook and supervised the project since January 1977.

The items in the *Number Test* have been prepared by Mr. Cornish, who undertook the work of conducting item-writing workshops, as well as the editing and reviewing of items.

It is hoped that mathematics teachers in Australian secondary schools will find this material of considerable value in assessing performance and diagnosing weaknesses among their students.

This handbook in the *ACER Mathematics Profile Series* is released as an experimental edition, and we invite comment on the material presented.

J. P. Keeves
Director
ACER

Contents

	Page
Foreword	iii
Contents	iv
List of Tables	v
List of Figures	vi
General Introduction	1
Purposes of evaluation using *ACER Mathematics Profile Series* Tests	2
Section I	
Number Test	
Introduction	4
Selection of an appropriate test	4
Instructions for test administration	
Before testing	5
Testing	5
After testing: interpretation of results	
Scoring	6
Converting raw scores to the MAPS scale	6
Ability interpretations	6
Visual interpretations using *Item Content/Process Profiles*	7
Diagnostic interpretations for individual students	8
Broader interpretations on the MAPS scale	9
Section II	
The Rasch measurement approach	
Background to the development of the MAPS scale	10
Advantage of the Rasch approach	10
The nature of the MAPS measurement scale	10
More details on relative mastery interpretations	11
Teacher-made tests and 'best test' design	13
Principals for teacher-made tests	13
Guidelines for 'best test' design	13
Standard error of measurement	15
Conclusion	15
Section III	
Broader interpretations in the *ACER Mathematics Profile Series*	16
Student development in terms of operational levels	16
Operational levels related to the MAPS measurement scale	16
Early concrete operational stage	17
Late concrete operational stage	18
Early formal operational (concrete generalization) stage	18
Late formal operational stage	20
Age/school year characteristics related to the MAPS measurement scale	20
Conclusions	21
Section IV	
Results and Conversion Tables	27
Bibliography and References	31

List of Tables

Tables		Page
1	Mean item *difficulty* and year range of *Number Test*	4
2	Odds of correct response and probability of success (W-scale and WITS scale)	11
3	Relative mastery levels (%)	12
4	Probability of success for integral differences between student *ability* and item *difficulty*	12
5	Standard errors of measurement (30 item 'best test')	15
6	Mean *difficulty* of items representing operational stages	17
7	Criterion *abilities* (in brytes) for various mastery levels in operational stages	17
8	Mean *MAPS ability* of students for ages 7 to 17 years	21
9	Process objectives	22
10	*Number Test:* Item *MAPS difficulties*	27

List of Figures

Figure		Page
1	Student S and five items located on the MAPS scale	1
2	Relationship of items and students on the MAPS scale	2
3	Mastery levels	2[
4	Answer and Record Sheet – recording MAPS *ability*	6
5	Answer and Record Sheet – relative mastery levels	7
6	Use of mastery profile 'cursor' against the *Number item process profile*	8
7	Relative *abilities* on 'best test' of various ranges	14
8	Relative mastery on operational stages and approximate age/school characteristics	24
9	*Number Test – item profile* – Process Areas	25
10	*Number Test – item profile* – Content Areas	26

General introduction

The *ACER Mathematics Profile Series* has been designed to provide a flexible system for monitoring students' mathematical development throughout their schooling.

The original intention of the programme was to develop a series of tests suitable for students at junior secondary level, and most of the tests specially created for the *ACER Mathematics Profile Series* have their focus at that level. However the decision to use a measurement scale based on the Rasch approach[1], made it feasible to link the results on all the tests in the *Series*, as well as other ACER tests[2], to the same scale. This has made it possible to extend the use of the tests into both primary and upper secondary school.

The common scale for the *ACER Mathematics Profile Series* will be referred to as the MAPS scale. This scale provides a convenient measure of the latent variable which all the tests have in common: best described as mathematical development. An important feature of the Rasch approach is that it enables both students and items to be located on the same scale.

In the *ACER Mathematics Profile Series* items are located by their *MAPS difficulty* and students by their *MAPS ability*.

The items have been located in relative positions on the MAPS scale according to the degree of mathematical challenge they offer students. These item measures have been called *MAPS difficulties*.[3]

For each set of items combined into a test, the marks obtained by students reflect the different levels of mathematical development elicited by the items offered. These total test scores can be easily converted into student measures on the common scale, thus giving the students' *MAPS abilities*.[4]

In the *Series*, *MAPS ability* and *MAPS difficulty* are defined according to this interaction between students and items.

Having a common scale provides the opportunity for a wide range of interpretations. The following example illustrates some of the potential afforded by the common scale.

The *MAPS ability* of a student S and the *MAPS difficulty* of five items are represented on the common scale in Figure 1. The bold numbers on the right of the scale represent the location of the items.

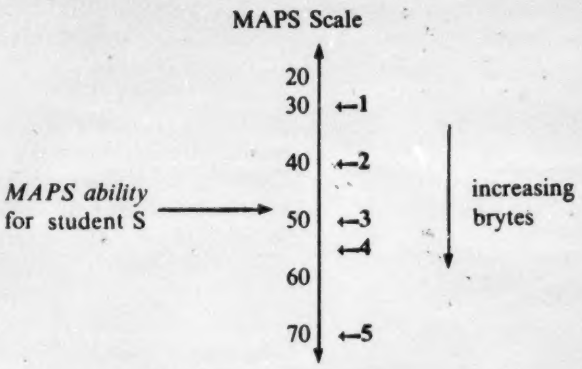


Figure 1: student S and five items located on the MAPS scale

The MAPS scale is marked in equal units called 'brytes'[5]; increasing 'brytes' indicates increasing *MAPS ability* of students (increasing mathematical development) and increasing *MAPS difficulty* of items.

Assume that student S has done a test and his test score has been converted to a *MAPS ability* of 48 brytes. Student S is then located on the scale by his *MAPS ability*.

An item is located on the MAPS scale by its *MAPS difficulty* also given in brytes. For example, item 3 has a *MAPS difficulty* of about 50 brytes. The five items are located on the MAPS scale using their respective *MAPS difficulties*. Items 4 and 5 are more difficult than item 3 and items 1 and 2 are easier.

As a consequence of the common scale certain interpretations are possible, for example:

(a) The *MAPS difficulty* of item 3 and the *MAPS ability* of student S are almost equal. This suggests that the student has roughly an even chance of getting item 3 correct.

(b) Item 1 would be relatively easy for student S, and item 5 would be relatively hard. This suggests that the student has a better than even chance of responding correctly to item 1, and a less than even chance of responding correctly to item 5.

(c) In selecting three items for student S, items 2, 3 and 4 may be considered to provide a suitable range: item 2 being slightly easy, item 3 moderate, and item 4 slightly hard.

This example displays some special interpretative features that are possible because of the MAPS scale. The most important of these is to suggest appropriate learning activities in terms of the students cognitive levels as well as their likely mastery of specific mathematical concepts and processes. This is possible because

(i) the scale has been broadly related to the Piagetian cognitive stages of development[6], and

(ii) all items in the *Series* have been classified according to content area and process objective[7].

[1] The characteristics and advantages of the Rasch approach to test measurement are discussed in Section II of this Handbook.

[2] For example, the ACER Primary Mathematics Tests (AM Series).

[3] *MAPS difficulties* were estimated from student responses using a computer program based on the Rasch approach. The *MAPS difficulties* for all test items are recorded in Section IV of the appropriate Handbook.

[4] Conversion tables from test scores to *MAPS abilities* are contained in Section IV of the appropriate Handbook.

[5] Wright and Douglas (1975) suggest the term 'brytes' as a convenient metric for measurement.

[6] Refer to Section III of this Handbook.

[7] Refer to *item profiles*, pp 25-26 and the foldouts in this Handbook.

The MAPS scale was arbitrarily centred on 50 brytes as this was considered to represent a level of mathematical development generally achieved at about the primary/secondary transition. Computer programs based on the Rasch approach make it possible to calibrate all item difficulties relative to this arbitrary point and hence to determine *MAPS abilities* on any sub-set of items whose *MAPS difficulties* are known. On this arbitrary scale student *MAPS abilities* on tests are linked to the MAPS scale range between 20 brytes and 80 brytes. Generally, tests suitable for primary students will give results which are converted to *MAPS abilities* in the range 25 to 50+ brytes, while those for junior secondary students lead to *MAPS abilities* in the range of 35 to 70+ brytes.

In theory, any student can be tested using any test, but the most accurate measures occur when the students answer a test suited to their level of mathematical development.

Purposes of evaluation using *ACER Mathematics Profile Series* tests

Usually tests are designed with a specific purpose in mind. These purposes include achievement testing (for ranking or grading), criterion-referenced testing (to ascertain whether students have satisfactorily mastered specified objectives), and diagnostic testing (to identify factors limiting student progress in certain areas). Often the nature of the test is predetermined by the purpose in mind, and the interpretation of results is limited to that purpose unless teachers have a sophisticated understanding of the measurement principles involved.

The properties of the MAPS scale enable multiple interpretations which may be useful for these purposes. The main focus of test development has been to select items which best fit onto the MAPS scale, so that this flexibility could be assured.

(i) Achievement rankings and diagnostic profiles

The common scale for all tests makes it possible to compare student *MAPS abilities*, whether derived from responses to the same test or to different tests. If desired, achievement ranking or grading across a group of students is possible. An alternative and probably more desirable purpose would be to draw up profiles for each student in a class. This could be done for different content areas and assist in determining the relative strengths and weaknesses of a student. Profiles across time are also valuable in a diagnostic sense; if a student is progressing more slowly than others in his class, the factors limiting his development may be identified by an investigation of items in the region of the scale near his present *MAPS ability*.

(ii) Criterion-referenced evaluation and relative mastery

The unique feature of the Rasch measurement approach is that it is possible to assess the likelihood of any student with known *MAPS ability* being able to correctly answer any item of known *MAPS difficulty*. This is represented in Figure 2 below.

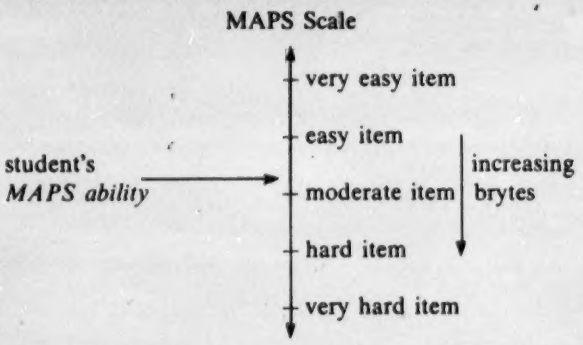


Figure 2: Relationship of items and students on the MAPS scale

Qualitatively, a given student has a better than even chance of succeeding on items with *MAPS difficulties* below his *MAPS ability* and a less than even chance of succeeding on items with *MAPS difficulties* above his *MAPS ability*. A student has an even chance of success on an item of the same *MAPS difficulty* as his *MAPS ability*.

Mastery is usually quoted in terms of percentage success over a group of items representing the desired objective. Criterion levels of 100%, 90%, and sometimes 75% are usually specified. Where student *MAPS ability* equals item *MAPS difficulty*, the even chance of success can be interpreted as 50% mastery. Quantitatively, the units on the MAPS scale have been chosen such that the usually quoted levels of mastery can be easily interpreted.

The scale in Figure 3 indicates the level of mastery for points on the MAPS scale where *MAPS ability* differs from *MAPS difficulty* by multiples of 5 brytes.

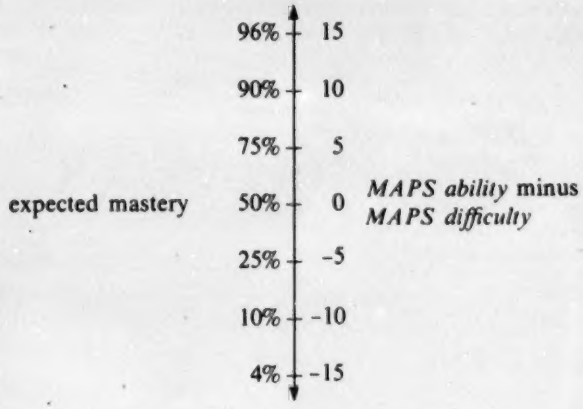


Figure 3: Mastery levels

Because the mastery levels are able to be related to any *MAPS ability* they are described as relative mastery levels. Depending on the task, it is possible to interpret these relative mastery levels in various ways. For example, suppose a teacher requires a student to have at least 90 % mastery in relation to items at 42 brytes. For 90% mastery, Figure 3 shows that the *MAPS ability*

should exceed *MAPS difficulty* by 10 brytes, so the student's *MAPS ability* should then be equal to or greater than 52 brytes (that is, 41 + 10 brytes). Alternatively, if a student of *MAPS ability* 32 brytes attempts items at 42 brytes (that is, 10 brytes above *MAPS ability*), the student would be expected to have 10% mastery of the items offered.

The potential for interpretation at a variety of levels makes it feasible to assess a student's status in relation to a sequence of learning objectives and to identify those objectives for which continued learning activities are likely to be of most benefit to a student. Theory suggests that the most suitable learning activities are likely to be reflected by items having *MAPS difficulties* within the range −5 to +5 brytes of each student's *MAPS ability*. The easier extreme (−5 brytes) should provide consolidation and confidence of objectives currently mastered to the extent of 75% and the more difficult extreme (+5 brytes) should provide a reasonable challenge with more than 25% likely mastery.

(iii) Cognitive levels and their implications for student learning

The interpretative features above are further discussed in this Handbook in relation to the specific tests used (Section I) and in relation to the Rasch measurement model (Section II).

Through the *Operations Test*[8] which is based on recent research on cognitive operational levels in mathematics (Collis, 1975), it has been possible to relate the MAPS measurement scale approximately to the Piagetian stages. It happens that the 25% to 90% mastery ranges at each cognitive stage follow the easily remembered pattern below.

Cognitive stage	25% to 90% mastery
Early concrete operational	30 to 45 brytes
Late concrete operational	40 to 55 brytes
Concrete generalization (early formal)	50 to 65 brytes
Late formal operational	60 to 75 brytes

Descriptions of the characteristic thought processes in mathematics for each of these stages are provided in Section III of this Handbook, together with items that reflect learning experiences appropriate during those stages.

A set of mathematics process objectives which relate to concrete operational and formal operation study of the subject are also included in Section III. The main subdivisions are arranged hierarchically, e.g. Knowledge, Operations, Relevance, Communication, Problem Solving, etc. and are used in the content/process grids for tests in the *ACER Mathematics Profile Series* other than the *Operations Test*. They may also provide additional ideas for learning experiences appropriate for students at each level.

Approximate age/year of schooling 'norms' are also recorded in Section III, but these may vary according to locality and mathematical background of students. It should be noted that the age/year of schooling means quoted at senior secondary level do not relate to all students, but only to those still involved in some mathematical study. The MAPS scale may be regarded as giving measures of mathematical development, and the age/year of schooling means and operational levels data should be interpreted in the same context.

[8] The *ACER Mathematics Profile Series—Operations Test* was the first test in the series to be published.

Section I
Introduction

When 'the man in the street' is asked the most important aspect of Mathematics courses, he usually responds in terms of skills of numerical manipulation. While teachers appreciate that Mathematics includes many systems besides the number system, most would agree that the development of number concepts is fundamental to broader mathematical understanding.

In the primary school years students are introduced to natural numbers and investigate interesting properties and patterns, such as adding successive 'odd' numbers to give 'square' numbers. In the junior secondary school, students begin to generalize relationships which underlie numerical manipulations, such as commutativity and distributivity and extend their investigations into the set of integers. They also begin to generalize fractional concepts and are able to investigate properties in the set of rational numbers. Eventually important concepts related to the real number system, such as identity and inverse elements, can be generalized. These generalized concepts and relationships are the basis of formal (algebraic) understandings of number system.

The *ACER Mathematics Profile Series – Number Test* consists of items designed to monitor students' development of number concepts, properties and relationships. In conjunction with the *Operations Test*, which focuses on the application of these relationships to the solution of increasingly sophisticated equations, the *Number Test* provides for diagnosis of student understanding of concepts which underlie confident manipulations in numerical problems and those involving algebraic formulae.

Many items for the *Number Test* have been written by teachers in various Australian states and represent a cross-section of experiences in the general area of number.

There are three units in the *Number Test* and each unit was designed primarily for use at the junior secondary level, although after examination of the three units teachers may decide that a unit is appropriate at the late primary level or the late secondary level. At the junior secondary level, the three *Number Test* units (Units I, II, III) increase in mean item *difficulty*, as shown in Table 1.

Each unit in the *ACER Mathematics Profile Series – Number Test* has been constructed to cover a variety of content areas and process objectives. These classified items are displayed in the *item content profile* and the *item process profile* which respectively display the four broad content groupings and three major process areas. The items are scaled within those categories against the MAPS measurement scale and are presented in the foldout sections on the front and back cover and pp. 25–26. Teachers would be familiar with the details of the content areas and the process groupings. Basic, Understanding and Application are common to the tests of the *ACER Mathematics Profile Series* and are detailed in Section III of the Handbook.

Selection of an appropriate test

Each of the three units in the *Number Test* covers a range of content and processes and these can be identified by referring to the *Number Test item profiles* (pp. 25–26 and the foldouts).

The *Number Test* units increase in mean item *difficulty* and are likely to be suitable for students in the school years as shown in Table 1.

Table 1
Mean item *difficulty* and year range of *Number Test*

Number Test	Mean item *difficulty*	Suggested year range
Unit I	50.8	7–8
Unit II	54.0	8–9
Unit III	57.4	9–10

However, because of the variety of experiences students receive in this area at different stages in their schooling, a definite prescription cannot be made. It is recommended that teachers become familiar with the items in the *Number Test* units before selecting an appropriate test.

To facilitate this, the three units in the *Number Test* are supplied in a single booklet, with a cover and directions page for each unit. Therefore it is easier to administer different units to different students in a class during the one testing session. This is permissible because the results on any unit (or any 'test' in the *ACER Mathematics Profile Series*) can be easily converted to the common MAPS scale.

Theoretically it does not matter which unit is given to students, but it is recommended that a unit be chosen so that most of the students are likely to respond correctly to one half or more of the items offered. For more able students at each year level, it may be appropriate to select a more difficult unit; for less able students, an easier unit. Because the units are in a single test booklet, administration in such situations is quite convenient.

An important advantage of the *ACER Mathematics Profile Series* is the potential to create 'new tests'. Because of the technique used to develop the *Series*, it is possible for teachers to select items from the tests in the *Series* and create 'new tests'. This flexibility, and the necessary guidelines for producing a 'new test' are explained in Section III of this Handbook. Discussions in Section III give information regarding approximate norms for age/year of schooling and operational levels. These may be used in conjunction with *Number Test item profiles* to select the most suitable unit.

Instructions for test administration

Before testing

1 Prior to administration teachers should become acquainted with the information provided in this Handbook and with the relevant set of items. This will assist in selecting the most appropriate unit and facilitate the interpretations.

2 The *ACER Mathematics Profile Series – Number Test* consists of three units, denoted *Number Test* Unit I, II and III. Each unit consists of 30 items and requires about 30 minutes actual testing time. An additional 10 minutes should be allowed for introductory procedures. The testing time has been suggested to suit the average duration of a class period. More time may be taken for testing if the teacher feels that it is necessary.

3 The practice questions, on the cover page of each unit, are common to the three units. Although the content of the practice questions comes under the topic of number, their prime function is to introduce students to the method of answering. The directions for administration are also common to the three units in the *Number Test*. This provides teachers with the option, if considered necessary, to administer more than one of the units in a session.

4 During the administration, the section enclosed in boxes are to be read aloud. These are suggestions to provide clear instructions; teachers may adapt them to better suit their particular students.

5 Students should have working paper available so that test booklets are not spoiled for subsequent use.

Testing

1 Direct students to fill in the section of the answer and record sheet with name, class, test, testing date, and birth date.

2 Direct students to the cover page of the appropriate unit. Inside the front cover of the *Number Test* booklet is a Table of Contents with the page number for the start of each unit. Depending on which is more appropriate for your particular class, either

(a) ask students to silently read the directions on the cover page, or

(b) ask students to follow silently as you read the directions aloud.

3 See the numerals I, II, III, on the top right-hand corner of your answer and record sheet? Circle the I (or whichever is appropriate) to show which unit in the test booklet you will do, then put your pencils down.

4 The purpose of the practice questions is to ensure that students understand how to record answers to the multiple choice questions.

Look at practice question **P1**. It says 'The measure of the shaded region in this diagram is called the Since the measure is called the area, the corresponding letter **A** has been marked on your answer and record sheet beside **P1**.

Students can be referred to their answer and record sheets to see how the answer to **P1** is recorded.

5 Now try practice question **P2** yourselves, and mark your answer and record sheet.

After about 30 seconds obtain the correct answer from a student

6 Yes, **C** is the correct answer so you should have marked the letter **C** on your answer and record sheet beside **P2**.

7 Are there any questions about this method of answering?

Answer any questions concerned with the method of answering.

8 You will have 30 minutes (or whatever time is appropriate) for this unit. Work as carefully as possible but do not spend too much time on any one question. If you cannot answer a question after thinking about it, leave it and go on to the next one. You may have time to come back when you have finished the rest of the unit. If you finish before time, check your answers, then close your booklet and wait quietly. Are there any questions?

If no further questions, say:

9 Turn over to the next page and start work now.

Record the time and check that all students have turned to the correct page.

10 After 30 minutes (or whatever time is appropriate), say:

Finish the question you are working on, then close your booklets.

11 As answer and record sheets are collected, check that the information is correct and entered clearly.

12 Check that all test booklets and answer record sheets have been handed in before concluding the testing activity.

After testing: interpretation of results

Scoring

The *Number Test* score key displays the correct responses for the three units in the *Number Test*. Answer and record sheets may be hand-scored by lining up the appropriate edge of the key with the question numbers and marking each answer. The key card displays the answers to Units I and II on the front face and Unit III on the back face. The back face of the score key card displays a mastery profile 'cursor' used in the identification of items as discussed in the sub-section titled **'Visual Interpretations using *Item Content/Process Profiles*'** (p. 7).

Referring to a particular unit, the total number of correct responses gained by each student should be recorded in the space designated *Raw Score* ... on the answer and record sheet. It would probably be most efficient to record all students' raw scores before converting these to the MAPS scale.

Converting raw scores to the MAPS scale

To convert raw scores to the MAPS scale, one should refer to the page of the results sections of this Handbook (Section IV) which is appropriate for the test given (e.g. page 28 for *Number Test* (Unit 1). At the top of this page, the mean and standard deviation of the *MAPS difficulties* for that set of items are recorded. The conversion of raw score to *MAPS ability* is presented in two ways.

(i) On the left hand side of the page, raw scores and associated *MAPS abilities* (correct to the nearest bryte) are listed, together with their associated scaled errors.

(ii) On the right hand side of the page, raw scores are presented against the appropriate points on the MAPS scale. As the *item content profile* and *item process profile* sheets are on this scale (pages 25–26), direct comparisons are possible.

Note that students who obtain zero or maximum marks cannot be rated on the MAPS scale using that test and should be required to answer an alternative test, either easier or harder respectively.

Using the left hand table, each student's raw score is converted to the MAPS scale and inserted in the middle box on the right hand top corner of the answer and record sheet, as shown in Figure 4.

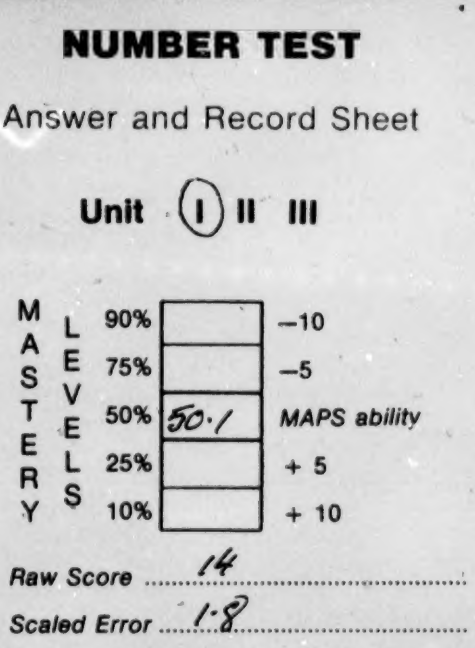


Figure 4: Answer and Record Sheet – recording *MAPS ability*

For example, a raw score of 14 out of 30 on *Number Test* Unit I is equivalent to a *MAPS ability* of 50.1 brytes, with an associated scaled error of 1.8 brytes (see Section IV, p. 27). The term 'brytes' was not written on the answer and record sheet to avoid student queries. A scaled error is reported for every *MAPS ability*. In the example given, a student's raw score of 14 corresponds to a *MAPS ability* of 50.1 brytes with an associated scaled error of 1.8 brytes. This scaled error can be used to suggest ranges in which it is likely that the student's mathematical development as assessed on the MAPS scale will lie. The range given by 50.1 ± 1.8 brytes is 48.3 to 51.9 brytes.[2] A more probable range in which the student's *MAPS ability* is likely to lie is given by 50.1 ± 2(1.8) brytes, that is 46.5 to 53.7 brytes.[3]

Ability interpretations

The role of educational evaluation is justified (in philosophical terms) if it incorporates interpretations which are both purposeful and meaningful to teachers and beneficial to their students. Tests in the *ACER Mathematics Profile Series* have not been designed for the ranking of students. Comparisons between students are recommended only if their *MAPS abilities* differ by more than the sum of the scaled errors associated with those *MAPS abilities*.

2 This range is usually defined as the 68% confidence interval.
3 This range is usually defined as the 95% confidence interval.

Comparisons of an individual's *MAPS abilities* at different points in time should also take due regard of measurement errors. From the above example, *MAPS abilities* of 50.1 and 52.1 brytes gained from successive testing do not necessarily imply improvement. The error associated with the student's test score could account for the difference in *MAPS ability*.

One major purpose of tests in the *ACER Mathematics Profile Series* is to identify the range of learning experiences which are likely to be beneficial for students at similar levels of mathematical development as assessed on the MAPS scale. This is possible, in quantitative terms, because of the specially chosen 'brytes' units of the MAPS scale.[4]

As mentioned in the general introduction, a student is likely to have 50% mastery over items with *MAPS difficulties* at the same point on the MAPS scale as his *MAPS ability*. On the student's answer and record sheet, the other four boxes can be easily filled in to give a profile of mastery levels (90%, 75%, 25%, 10%) relative to certain points on the MAPS scale. These points are calculated by adding −10, −5, +5, +10 respectively to the recorded *MAPS ability*, as indicated in Figure 5.

NUMBER TEST

Answer and Record Sheet

Unit (I) II III

MASTERY LEVELS			
90%	40·1	−10	
75%	45·1	−5	
50%	50·1	*MAPS ability*	
25%	55·1	+ 5	
10%	60·1	+ 10	

Raw Score ..... 14 .....

Scaled Error ..... 1·8 .....

Figure 5: Answer and Record Sheet – relative mastery levels

This implies that a student with a *MAPS ability* of 50.1 brytes would find items at 40.1 brytes very easy and items at 60.1 brytes very hard, since

(i) 40.1 brytes corresponds to 90% mastery or odds of success 9:1, and

(ii) 60.1 brytes corresponds to 10% mastery or odds of success 1:9.

In general, most students should benefit from learning experiences exemplified by items in the range −5 brytes and +5 brytes about their *MAPS ability*. This corresponds to a range of mastery from 25% to 75% respectively. So most students with *MAPS ability* of 50.1 brytes should benefit from learning experiences exemplified by items in the range 45.1 to 55.1 brytes.

Within this suggested range, the student's odds of success would range from 3:1 for items at 45.1 brytes down to 1:3 for items at 55.1 brytes. The easier items in this range, items between 45.1 and 50.1 brytes would represent learning experiences with which the students show some familiarity, but consolidation and practice would promote greater confidence. Items between 50.1 and 55.1 brytes would be answered with lesser confidence (50% to 25% mastery). These suggest learning experiences for which assistance such as teacher explanation and the use of concrete aids should be available during the learning process.

Teacher awareness of students' personality might also suggest modifications to the most suitable learning range. Less confident students should be offered a majority of experiences in their 50% to 90% mastery range, but those with high motivation may react well to occasional challenges where their likelihood of success is less than 25%.

Visual interpretations using *Item Content/Process Profiles*

The items in the *Number Test* have been classified according to their content areas and process objectives. This information is presented visually in the two fold-outs and pages 25–26 which are labelled *item process profile* and *item content profile*.

The *item profiles* serve two main purposes:–

(i) they display the relative item difficulties across the three units, against the MAPS measurement scale, so that comparisons can be made with broad content and process categories.

(ii) they facilitate teacher interpretations of students' likely mastery on all items, not only for the items answered.

The mastery profile 'cursor' (displayed on the reverse of the score key card) enables teachers to easily identify those items which lie in the mastery ranges discussed above. By placing the arrow ➡ against the student's *MAPS ability*, the items for which the student is likely to show 90%, 75%, 50%, 25%, and 10% mastery can be visually scanned. The technique is shown in Figure 6.

4 Reasons for this potential of the MAPS scale are discussed in Section II of this Handbook.

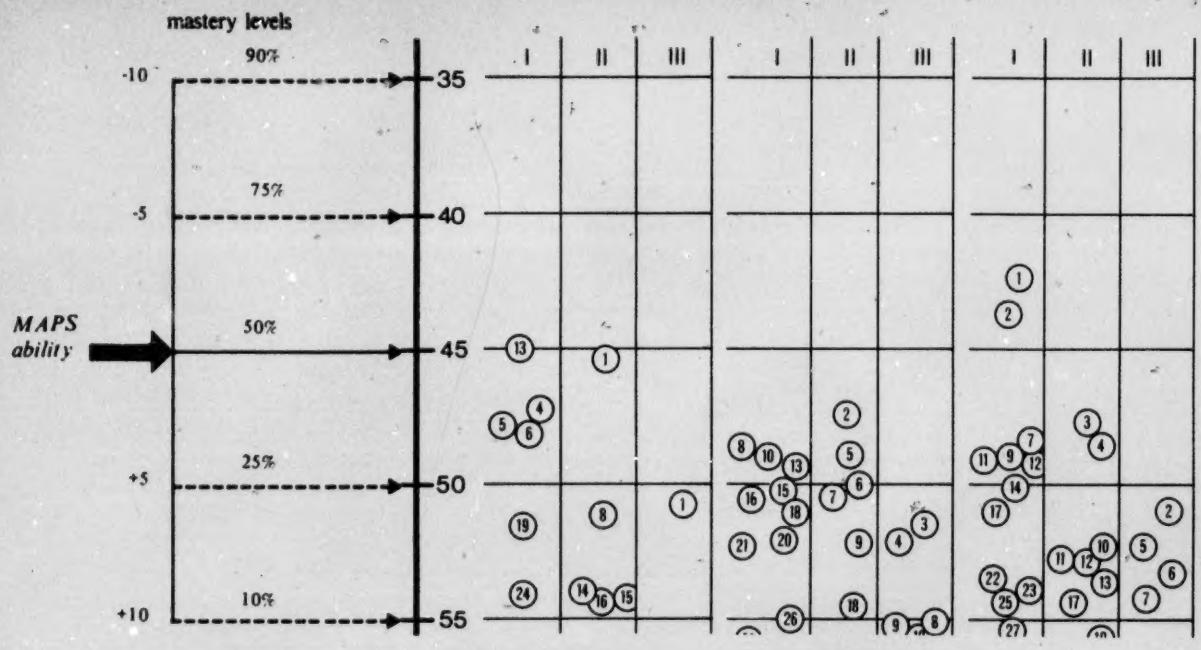


Figure 6: Use of mastery profile 'cursor' against the *Number item process profile*

The likely mastery of intermediate items can be interpolated visually or may be calculated more precisely using Table 4 (Probability of success for integral differences of student *ability* less item *difficulty*) which is provided in Section II of this Handbook. However, considering measurement errors, the mastery levels discussed above should be sufficiently precise for most interpretations.

Diagnostic interpretations for individual students

The units in the *Number Test* have been structured so that the items form a relatively even spread over the four content areas and three process objectives (presented in the foldout section).

If the items which a student has answered incorrectly are identified on his answer and record sheet, it is then beneficial to observe the types of items for which those incorrect responses were made. This should be in terms of both the content and process classifications. For example, while a student is responding correctly to 'Natural number' items at the Basic process level, he may be finding difficulty with items at the Application process level, in the same content area.

This diagnosis should mainly focus on items around the student's *MAPS ability* (i.e. where the odds of a correct response are approximately 1:1). Such items could suggest to teachers learning experiences that are most likely to be of use in remedial work. The actual error or misconception may be identified by looking at the incorrect option chosen by a student and considering possible reasons for the student's selection. As mentioned previously the item classifications may also provide useful information. Teachers may also find a very useful way to gain information is to ask students why they selected a particular incorrect option. Common misconceptions may be unearthed as the result of such questions, and suitable remediation strategies may be developed.

Broader interpretations on the MAPS scale

Other tests in the *ACER Mathematics Profile Series* (e.g. *Operations Test*) and some of the *ACER Mathematics Tests* (*AM Series*) have been linked to the same MAPS scale. For this reason student *MAPS abilities* based on the units in the *Number Test* may be useful to predict likely student success in these other areas. While consistent development across the different topic areas of mathematics is not assumed, such interpretations may provide a reasonable estimate or starting point until more precise results from testing in those specific areas are obtained. The same visual procedures (relative mastery profile 'cursors' laid adjacent to the appropriate *item profiles*) would apply, with identification of suitable learning experiences by perusal of the relevant test booklets. The Handbooks for those tests would supply more specific information on item difficulties and content/process classifications.

The *ACER Mathematics Profile Series* Handbooks also contain information on item difficulties and score conversions to the MAPS scale for some ACER test materials which may be used in conjunction with each series [e.g. *Jenkins Non-Verbal Test* in *Operations* Handbook, *AM Series spatial tests* (*AM8, AM9, AM10, AM12*) in *Space* Handbook]. These tests should be administered as recommended by their respective Handbooks, although some extra time may be allowed if used diagnostically.[5] Since item ***difficulties*** are provided, relative mastery interpretations may suggest additional learning activities, even if the tests are not administered. The potential to make use of a wide variety of items from many tests, although the tests may not have been administered, expands the pool of items available to teachers for selection.

5 In this situation the normative data in the Handbooks cannot be used.

Section II

The Rasch measurement approach

Background to the development of the MAPS scale

The MAPS scale is based on a probabilistic measurement model introduced by George Rasch of the University of Copenhagen in the early 1960s. Procedures to apply the model for educational testing purposes were developed by Benjamin Wright of the University of Chicago during the late 1960s.

Advantages of the Rasch approach

The main advantage of the Rasch measurement model is that it provides *objectivity* of measurement. Rasch (1966) has discussed this concept in mathematical terms but Wright (1968) has considered the *objectivity* of the Rasch approach in practical terms, using provocative contrasts with traditional procedures. He emphasizes the potential of the Rasch approach for *sample-free test calibration* compared with conventional tests statistics based on student samples which are representative of particular populations specified by age, year of schooling, and/or geographical locality. Wright also advocates the benefit of *item-free person measurement* whereby students can answer different sets of items but their abilities can be measured on the same scale (conventionally only possible with parallel test forms).

The Rasch approach makes it possible to provide a common measurement scale on which student ability and item difficulty are simultaneously calibrated, independently of any external characteristics of the students tested or the items answered, except those relevant to their interaction (i.e. to the latent trait being analysed). However once the scale is established it is possible to superimpose information concerning relevant characteristics of students or items onto the scale. For example, the ages or years of schooling for any desired population of students can be superimposed by analysing[1] the abilities displayed by representative student samples. Similarly it is possible to categorize items according to certain characteristics and superimpose those categories onto the scale by analysing the difficulties of representative samples of items.

Hence the Rasch measurement approach preserves the potential for 'normative' student information, but offers this in the context of objective measurement (independent of population statistics), and gives the bonus of the potential for 'normative' item information. In the *ACER Mathematics Profile Series* the latter potential has been used to superimpose onto the established scale the Piagetian levels of mathematical development. So, in essence, it is the established scale which is being 'normed' rather than the test.

[1] 'analysing' in this context implies finding the distribution of student abilities or item difficulties within each category of the chosen characteristic. For example, for a representative sample of students, the means (and dispersions) of MAPS abilities in each age category could be calculated and related to the MAPS scale. The overall probability of students of a given age mastering items on the scale can be assessed.

Another major advantage of the Rasch measurement scale is that its probabilistic basis allows for interpretations in terms of odds of success or relative mastery over all items calibrated onto the same scale, not only those actually attempted by the student, to determine his ability. The first tests to make use of this potential, the KeyMath Diagnostic Arithmetic Test and the Woodcock Reading Mastery Tests, were published by American Guidance Services and are available in Australia through ACER.[2] Both of these tests provide student profiles across different aspects of the ability measured, since estimates of ability on these different sub-scales are regarded as comparable with the original scale in each case.

The advantages of the Rasch approach (objectivity of measurement, 'normative' potential, interpretation in terms of relative mastery, and profile comparability) are clearly relevant to the development of series of associated tests, such as the *ACER Mathematics Profile Series* and the AGS tests, for use by individual teachers in circumstances they choose. However the Rasch measurement approach has also been applied in a broader context across large educational systems. For example, at the National Foundation for Educational Research in England, interest has focussed on its usefulness for national survey testing and for item banking designed to aid consistency in school-based certification (Sumner, 1975). In the United States, similar uses have been initiated by regional boards of education (Forbes and Ingebo, 1975). The Rasch approach is also being used in education research throughout the world to investigate the provision of consistent and reproducible scales for both cognitive and attitudinal traits.

The nature of the MAPS measurement scale

The Rasch model of measurement of any trait is based on the simple idea that odds of success by a person on a particular task depend on an interaction between the ability of the person and the difficulty of the task. When the person's ability is greater than the task difficulty then the odds of success are high (3:1 say); when the person's ability is less than the task difficulty, then the odds of success are low (1:3 say). If a person's ability is the same as the task difficulty, then the odds of success are even (or 1:1), and the probability of mastery is 50%. For Rasch measurement scales, the concepts of ability and difficulty are defined in terms of this concept of relative odds of success which must be consistent across a variety of interactions between people and tasks.

Computers have made it possible to analyse the complicated interactions between people of differing abilities and items of varying difficulties. Relative odds of success are estimated and refined by iterative procedures until they stabilize.[3] Items which display inconsistent relative odds for persons of varying ability are rejected, and the estimation procedures are repeated without their influence. The necessity of consistent rela-

[2] KeyMath Diagnostic Arithmetic Test is stocked in Australia, and the Woodcock Reading Mastery Test is imported on request.

[3] For detailed information refer to Wright and Panchapakesan (1969).

tive odds of success has been clearly exemplified by Sumner (1975, p. 108).

The mathematics of the Rasch model and associated computerized procedures are facilitated by use of the exponential function:-

$$\text{odds of success} = e^{(\text{ability} - \text{difficulty})}$$

Consequently, the ability and difficulty estimates obtained are initially on a log-odds scale, where one unit difference between ability and difficulty represents odds of success of e:1 (or 2.718:1). Many research reports use this log-odds scale, which usually ranges from −4 to +4 units, centred on an arbitrary zero which represents the average log-odds difficulty of the items being considered.

The log-odds scale is convenient for theoretical purposes and satisfactory for research studies being read by those with statistical expertise. However for practical interpretations, it is somewhat awkward to work out probabilities associated with differences on the log-odds scale. For example,

if ability − difficulty = 2 log-odds units
then odds of success = e^2:1
= 2.718^2:1
= 7.3791:1.

Hence probability of success is

$$\left(\frac{\text{odds}}{1 + \text{odds}}\right) = \frac{7.3891}{1 + 7.3891} = \frac{7.3891}{8.3891} = .881$$

which is equivalent to 88.1% relative mastery.

To avoid such tedious calculations, it is more convenient to transform the log-odds scale to one where differences along the scale relate to the easily remembered probabilities discussed in the interpretations section of this Handbook (that is, .75, .90, etc.). The transformation most commonly used centres on 50 instead of zero, and is given by the formula

$$D = 50 + 4.551d$$

where D is the new scale value, and d is in log-odds units. The transformation factor 4.551 represents 5/log e^3 and this implies that an increase of 5 units on the new scale represents a growth factor of 3 in the odds of success.

The measurement scale has been discussed as the W-scale by Woodcock and Dahl (1971), and as the WITS scale by Wilmott and Fowles (1974). The latter publication, which summarizes early research at NFER on the application of the Rasch model, provides an interesting analysis of objectivity in measuring test performance compared with that in physical measurement. An appendix also provides a relatively simple version of the mathematical equations used in computer analyses for the Rasch model. Neither of these sources names the units of their scales, but the name 'brytes' has been used by Wright and Douglas (1975).

The focus of discussions on the W-scale and the WITS scale was the potential for consistent interpretations of odds of success and relative mastery probabilities for particular differences along their scales, as shown in Table 2.

Table 2
Odds of correct response and probability of success (W-scale and WITS scale)

Difference (ability − difficulty)	Odds of correct response	Probability of success
−20	1:81	.01
−15	1:27	.04
−10	1:9	.10
−5	1:3	.25
0	1:1	.50
+5	3:1	.75
+10	9:1	.90
+15	27:1	.96
+20	81:1	.99

The MAPS scale used in the *ACER Mathematics Profile Series* also displays these interpretative features. It differs from these other scales in that the arbitrary centre of 50 brytes represents a different level of ability from, say, the W-scale in the KeyMath Diagnostic Arithmetic Test which is centred about the middle of infants-primary schooling. For the MAPS scale, 50 brytes has been assigned to the *MAPS ability* level where students have an even chance of mastering the simplest items which require 'formal' operational thinking (see p. 18). The rationale for this decision is explained more fully in the following section. However, as mentioned in the general introduction, 50 brytes on the MAPS scale may be broadly interpreted as a transition point after which students may be ready to appreciate the structured system of mathematics.

The MAPS scale was initially calibrated using the items in the *Operations Test*, and all other items have been linked to this MAPS scale. Briefly, the procedures for linking involve using common items across trial forms within any area (e.g. *Space Test* units) and common students across areas (e.g. *Operations Test* and *Space Test* units). In all cases, the stability of the variance between difficulty estimates from both sources was determined to be statistically significant before linkage translations of scale were made. Correlation coefficients between linked item difficulty estimates were generally in the range .95 to .99 (mean r = .965, sd = .023 within *Space Test* units and .974 between *Space Test* units and *Operations Test* scales). These results suggest that, with due regard for the standard errors of measurement associated with any student's individual score, interpretations between any of the *ACER Mathematics Profile Series* tests using the MAPS scale can be made with confidence.

More details on relative mastery interpretations

It is more important to emphasize that relative mastery interpretations are consistent for any given difference between person *MAPS ability* and item *MAPS difficulty* on the Rasch measurement scale. The consistency and symmetry of this interaction is well displayed in Table 3.

Table 3

Relative mastery levels (%)

item \ person	*MAPS ability* (brytes)								
	35	40	45	50	55	60	65	70	75
35	50%	75	90	96	99	100	100	100	100
40	25	50%	75	90	96	99	100	100	100
45	10	25	50%	75	90	96	99	100	100
50	4	10	25	50%	75	90	96	99	100
55	1	4	10	25	50%	75	90	96	99
60	0	1	4	10	25	50%	75	90	96
65	0	0	1	4	10	25	50%	75	90
70	0	0	0	1	4	10	25	50%	75
75	0	0	0	0	1	4	10	25	50%

MAPS difficulty (brytes)

A relative mastery level such as 75% can be interpreted in various ways, depending on the situation. So far the emphasis has been placed on the concept of probability of success by an individual person succeeding on individual items that as yet have not been attempted. For example, a person with *MAPS ability* 60 brytes would have 75% chance of correctly answering an item of difficulty 55 brytes (odds of success = 3:1). Of course, when answering an item, a person is either right or wrong. However, if a person with *MAPS ability* 60 brytes answered a number of items with difficulty 55 brytes, he would be likely to answer three-quarters of them correctly. Alternatively, when a group of persons all with *MAPS ability* 60 brytes answer one item at 55 brytes, three-quarters of them are likely to be correct, this proportion being the traditional statistic, *item facility*. This idea can also be extended to groups of people answering groups of items evenly distributed about a particular difficulty level, which is the basis of 'best test' design discussed later in this section.

In Table 3 the *MAPS abilities* and *difficulties* were multiples of 5 brytes, but the relative mastery levels also apply for intermediate values. If a person with *MAPS ability* 48 brytes attempts an item with *difficulty* 53 brytes, the difference is −5 brytes, so the probability of success is 25% (odds of success 1:3). This feature is used on the answer and record sheets to identify the relative mastery levels for each student. For most purposes the conveniently remembered percentages for differences of multiples of 5 brytes are sufficient. Mastery probabilities associated with interpolated differences are not easily remembered and, when needed, are available from Table 4.

Table 4

Probability of success for integral differences between student *ability* and item *difficulty*

brytes	P*	brytes	P*	brytes	P*
25	.996	8	.853	−9	.122
24	.995	7	.823	−10	.100
23	.994	6	.789	−11	.082
22	.992	5	.750	−12	.067
21	.990	4	.707	−13	.054
20	.988	3	.659	−14	.044
19	.985	2	.608	−15	.036
18	.981	1	.555	−16	.029
17	.977	0	.500	−17	.023
16	.971	−1	.445	−18	.019
15	.964	−2	.392	−19	.015
14	.956	−3	.341	−20	.012
13	.946	−4	.293	−21	.010
12	.933	−5	.250	−22	.008
11	.918	−6	.211	−23	.006
10	.900	−7	.177	−24	.005
9	.878	−8	.147	−25	.004

*P = Probability of success

Relative Mastery percentages are given by P × 100.

The more specific interpretations using integral differences may be useful for precise diagnoses of individual students' difficulties through a learning sequence implied by a group of related items. For example, in a sequence represented by items at 42, 43,

47; 48 and 51 brytes, a student with *MAPS ability* 46 brytes would, from Table 4, have the following probabilities of success: 0.71, 0.66, 0.45, 0.39 and 0.25. If the student lacks confidence, it may be desirable to concentrate on consolidating activities at a level appropriate for the items at 42 and 43 brytes, leading up to experiences at a level appropriate for the items at 47 and 48 brytes.

Teacher-made tests and 'best test' design

Although the tests offered in the *ACER Mathematics Profile Series* provide considerable flexibility for teacher use, their basis on the Rasch measurement model makes it possible to create new 'tests' selected from items whose difficulties are calibrated onto the MAPS measurement scale. The background principles described below must be considered and simple but specific guidelines should be followed in order to create a blueprint for new tests which best suits the evaluative purposes.

Principles for teacher-made tests

The items in all areas of the *ACER Mathematics Profile Series* are calibrated onto the MAPS scale which aims to measure mathematical development across all content areas. The Rasch measurement model is sometimes criticized by measurement specialists in that it assumes that interactions depend on a single latent trait. However it is more appropriate to say that the trait is whatever all the items have in common, assuming that constituent factors such as process and stimulus occur randomly (and hence proportionally as they contribute to the trait).

Consequently the *first principle* in creating a new test is to choose items which reflect the general nature of the total set of items. To focus predominantly on one type of item (e.g. verbal knowledge in trigonometry) would affect the interpretative potential in relative mastery terms.

The *second principle* relates to the evaluative purpose which determines the blueprint of the test (see **Guidelines for 'best test' design** below). For survey and achievement purposes, the mean and range of the item difficulties that form a test should match the mean and range of likely student *MAPS abilities*, which may be inferred from previous testing or otherwise may be estimated using the approximate 'norms' provided in the next section. However when a test is being designed to divide students at a particular cut-off point (for example, to group students for different learning strategies) the mean of the test should be that cut-off point and the range should be as narrow as possible. Tests of less than 20 items in length provide less precise information for measurement (although they may serve motivational purposes), and the longer the test the more accurate will be the resultant interpretations.

A *third principle* links professional standards and the copyright of items in the *ACER Mathematics Profile Series*. Any creation of a new test based on these items should be carried out only by educators who thoroughly understand the principles of mathematics education and evaluation covered in this Handbook and who are competent to make the interpretative judgments implied in the remainder of the section. Fundamentally, new tests should be created more as a supplement to, or broad extension of, the *ACER Mathematics Profile Series* as presently offered by ACER.

Guidelines for 'best test' design

These guidelines are based on experimental work carried out by Benjamin Wright and Graham Douglas at the University of Chicago (Wright and Douglas, 1975). Their research has shown that the 'best test' design from items calibrated onto a Rasch measurement scale is a very simple design. This involves creating a test where the item difficulties are evenly distributed along the scale centred on a point M brytes, the mean of the test. This design has the added advantage that it is very easy to read off *MAPS abilities* relative to M brytes, for any test of given length (L items) and range (R brytes). The range of the test depends on the likely range of *MAPS ability* of the students. Survey tests would have a wide range (R = 15 brytes or more), but a test used to make fine discriminations about a cut-off point would require a narrower range of 5 or 10 brytes. A test with a narrow range is more 'sensitive' to differences in student *MAPS abilities*. Also the greater the length (L items) of the test, the lower the standard error of measurement associated with each test score.

Suppose it has been decided to create a 30 item test focusing on a mean of 55 brytes. Considering item difficulties to the nearest bryte, a survey test with a range of 15 brytes should consist of 2 items at each bryte from about 48 to 62 brytes. A test to divide students at 55 brytes may have a range of 5 brytes, so six items at each bryte from 53 to 57 brytes would be chosen. In each case the actual items chosen vary a little from the desired 'blueprint' for the 'best test', without seriously affecting the interpretation of the results. The value of the test mean for the chosen items should be calculated using the reported item *difficulties*, and the calculation involves finding the average value of the item *difficulties* (see Section IV). This is the final value from which student *MAPS abilities* should be calculated.

Results on the test are recorded as raw scores (out of 30 in the example above) then treated as the proportion correct. Figure 7 shows such proportions for tests where R = 5, 10, 15, 20, 30, and 40 brytes, scaled against *MAPS abilities* relative to the mean of the test. To obtain the *MAPS ability* of a student the relative *MAPS abilities* in brytes (on the vertical axis of Figure 7) are

added to the mean of the test if the proportion correct is greater than 0.50, and

subtracted from the mean of the test if the proportion correct is less than 0.50.

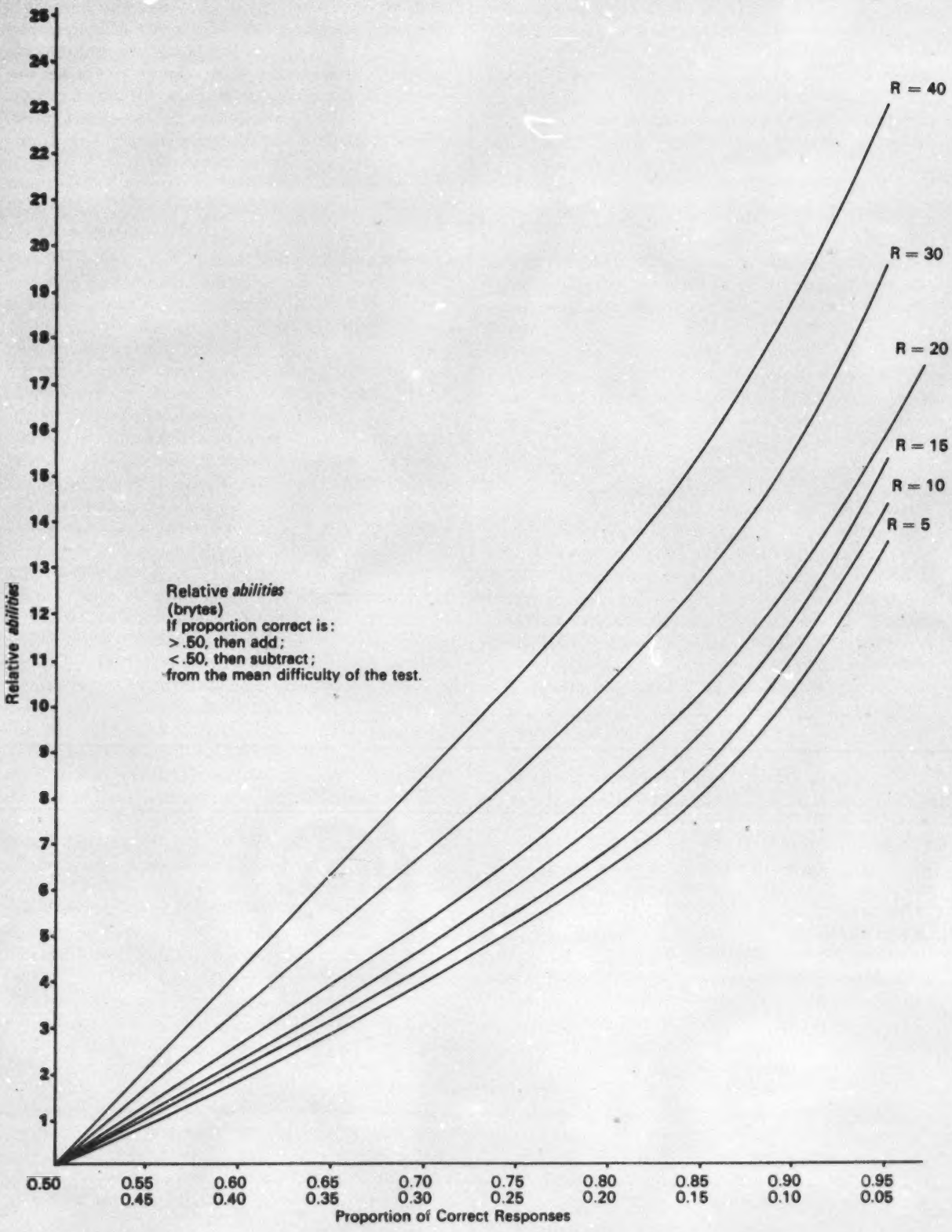


Figure 7
Relative *abilities* on 'best tests' of various ranges

For example, where M = 55.4 brytes and R = 15 brytes a score of 24 out of 30, or proportion correct 0.80, is equivalent to (55.4 + 7.5) brytes or 62.9 brytes. On the same test, a score of 10 out of 30 (proportion correct 0.33) is equivalent to (55.4 – 4.3) brytes or 51.1 brytes.

Standard error of measurement

Standard errors of measurement on 'best tests' vary according to the length, range, and proportion of correct responses. The values in Table 5 (given in brytes) are stated for a 30 item test. For the above example, 62.9 brytes can be represented as M + 7.5 brytes. From Table 5 when R = 15 brytes, the standard error of measurement is about 2.3 brytes. On the same test, 51.1 brytes can be represented as M – 5 brytes, so the standard error of measurement is 2.0 brytes. For any other test of length L items, they should be multiplied by a factor of $\sqrt{30 \div L}$.

Table 5

Standard errors of measurement (30 item 'best test')

Ability	R = 5	R = 10	R = 15	R = 20	R = 30	R = 40
M brytes	1.7	1.7	1.8	2.0	2.2	2.5
M ± 5 brytes	1.9	2.0	2.0	2.1	2.3	2.5
M ± 10 brytes	2.8	2.6	2.6	2.5	2.5	2.6
M ± 15 brytes	4.2	4.1	3.9	3.7	3.1	2.8
M ± 20 brytes	–	–	–	–	3.5	3.5

It should be noted that the minimum standard error of measurement occurs at the centre of a test with a narrow range, which is the reason that such tests should be used for dividing students into two groups. (This situation is equivalent to the criterion in traditional test statistics, that tests for pass/fail at 50% maximise the 'spread' of students if item facilities are close to 50%.) In both traditional and Rasch measurement, standard errors of measurement are decreased with longer tests. For example, the standard error of measurement at the centre of a narrow test (R = 5 brytes) of 50 items would be 1.3 brytes (i.e. $1.7 \times \sqrt{30/50}$); with 100 items the standard error of measurement would be reduced to 0.9 brytes.

Conclusion

The Rasch measurement scale offers both objectivity of measurement and flexibility of evaluation. An understanding of this section should provide those using the *ACER Mathematics Profile Series* with information with which they can utilize the full potential of the Rasch measurement approach. Use of the prepared tests, or those created according to the principle of 'best test' design, should cover most evaluative situations, affording maximum accuracy and interpretative potential. *Sample-free item calibration* has been the main focus for the test developers, but the duality of *item-free person measurement* is the main advantage for teachers. Measurements can be made using different 'tests' or sets of items chosen in each case as being the most suitable for both the students and the evaluative purposes. The implications for students' subsequent learning experiences are also an important potential and will be discussed as they apply to the MAPS scale in the subsequent section.

Section III

Broader interpretations in the *ACER Mathematics Profile Series*

– in terms of operational levels and other normative data

The MAPS measurement scale has been discussed in terms of its potential to identify suitable learning activities for each student, ranging from those which promote confidence (75% – 50% mastery) to those with motivational challenge (50% – 25% mastery). It is also possible to make broader interpretations concerning learning experiences or activities if the level of challenge they embody can be related to the Piagetian operational stages. This is feasible because the MAPS scale was originally calibrated using items similar to those which Collis (1972, 1975) used in his research on Piagetian levels in mathematics development. The factors he identified make it possible to superimpose on the MAPS scale regions which reflect student progress through the various stages of operational thinking.

Student development in terms of operational levels

The application of Piagetian theory can provide teachers with a valuable framework against which the appropriateness of learning activities can be considered. This becomes most important when teachers tailor educational experiences to their particular student groups rather than follow prescribed syllabi.

In a framework relevant to primary and secondary schooling, it is convenient, at least initially, to consider student development from the point of view of the two broad stages – concrete operational and formal operational. One feature of concrete operational thought is that the strategy needed to perform an operation needs to be represented, or mentally justified, by a corresponding manipulation of appropriate concrete materials. In contrast, formal operational thinkers can make use of hypothetico-deductive strategies, which involve reasoning on the basis of various alternative possibilities, to consider the interaction of relevant factors which may not be apparent in the concrete sense. Formal operational thinking can also be described as 'second-order' thinking – being able to operate on operations or relationships. Just as one criterion of concrete operational thought is the ability to organize objects and their familiar attributes into systems (by classification or seriation), formal thought implies an understanding of how operations relate together and a capability of rearranging relationships within such systems.

Most of the research in operational thinking has focused on the structured systems of mathematics and science. In the Australian Science Education Project, the teachers' guide provides comparisons between concrete and formal operational understanding in twelve characteristic areas (Australian Science Education Project, 1974, especially pp. 64–67). Many of these areas are also relevant to mathematical activities, for example: specifying position; understanding area and volume; dealing with probability and correlation; establishing proportional relationships between variables.

Similar criteria can be applied when assessing student reasoning in areas where the situations are not systematically structured. In their investigation of student judgment in the social and physical sciences, Peel and his colleagues (1972) have categorized responses at the following levels:

- logically restricted responses: mere repetition or rephrasing of information, ignoring asserted facts, introducing irrelevancies;
- circumstantial conceptualization: concrete-bound inferences, taking account of one or two most dominant aspects;
- logically possible responses: exploring some aspects hypothetically but unable to perceive overall effects;
- deductive conceptualization: hypotheses developed and analysed comprehensively.

Consistent responses in the first two categories would suggest early and late concrete operational thought, while the latter two suggest progress through the formal operational period.

The fact that these qualitative categories have been applied to a variety of eductional disciplines would suggest that it should be possible, and probably easier, to apply them across a wide range of different mathematical experiences. In conjunction with interpretations of the Rasch measurement approach, these qualitative criteria can also be quantitatively related in relative mastery terms. This avoids the dependence on relating operational levels to student age, a feature of Piagetian research which has caused some confusion and disagreement.

Operational levels related to the MAPS measurement scale

Items similar to those in the *Operations Test*, were offered in open response format to a wide range of students during Collis's early research on levels of reasoning in mathematical operations. Of the factors embodied in those items, he found that both the *structure* of the operation and the nature of the *elements* operated upon, influenced item difficulty. Even when correct responses were given, the appropriateness of reasons offered increased with overall score on the items. Item structure was broadly classified as concrete or formal, and this transition was found to be a more significant influence on item difficulty than increasing abstractness of the elements, that is, small to large numbers to pronumerals. Collis has summarized the transition as follows (Collis, 1975, p. 91):

> The evidence in this study, for operations (structure) being the prime cause of difficulty in items, is fairly clear but must be considered in relation to the elements involved also. The results seem to show that, given constant operational level, the abstractness of the elements determines the difficulty but, in addition, the operational difficulties are more readily overcome when the elements are simple. There is

clearly an interaction here and it is most noticeable between ten and fifteen years.

During this interactive phase, students demonstrate the ability to generalize familiar patterns (those whose structure can be related to concrete manipulations) in abstract terms, such as $(j - k) + k = \Delta$. They can also handle more formal structures such as $7 - 4 = \Delta - 7$ (not familiar from concrete manipulations as long as the elements are simple enough to be verified by numerical closure). However they would not be capable of explaining the logical rules for 'operating on operations' required in the parallel item, $f - e = \Delta - f$: the second order capability would be a necessary characteristic of truly formal operational thinking. Collis has termed this interactive level as the *concrete generalization stage*, to emphasize the restricted nature of students' capabilities, in preference to Piaget's term 'early formal operations'.

In calibrating the MAPS measurement scale, the onset of this *concrete generalization stage* has been chosen as an arbitrary point of reference to assist teacher interpretations, and the difficulty of the first items which represent this capacity has been translated to a value of 50 brytes. Furthermore, it is possible to analyze the scale according to the average difficulties of groups of items which display the characteristics of early and late concrete operations, and early (concrete generalization) and late formal operations. The results of the analyses are summarized to establish regions of the MAPS scale onto brytes. Furthermore, it is possible to analyse the scale superimposed. The characteristics of the representative items can then be discussed and generalized to broader activities.

Table 6

Mean *difficulty* of items representing operational stages

Operational stage	Characteristic items in *Operations Test*	Mean *difficulty* of items
Early concrete operational	1-7, 21-25, 41-43	35.0 brytes
Late concrete operational	8-14, 26-32, 44-50	44.6 brytes
Early formal concrete (generalized)	15-20, 33-36, 51-54	54.9 brytes
Late formal operational	37-40, 55-60	65.1 brytes

The average difficulties of these item groups increase approximately 10 brytes for each operational stage. Remembering that the MAPS measurement scale can be interpreted simultaneously with respect to student *MAPS ability* or item *difficulty*, it is reasonable to infer that students would have 50% mastery within each stage when their *MAPS abilities* correspond to the mean difficulty levels reported in Table 6. It is also interesting to consider points at which students are likely to progress through 25%, 75%, and 90% mastery of a range of tasks typical of each stage. For convenience, the *MAPS abilities* have been stated to the nearest bryte.

Table 7

Criterion *abilities* values (in brytes) for various mastery levels in operational stages

Operational stage	Mastery 25%	50%	75%	90%
Early concrete operational	30	**35**	40	45
Late concrete operational	40	**45**	50	55
Early formal (concrete generalized)	50	**55**	60	65
Late formal operational	60	**65**	70	75

The results indicate that 75% mastery at any one stage is equivalent to 25% mastery at the next. The overlap can be clearly observed in Figure 8 (on page 24) which summarizes all normative data against the MAPS measurement scale. For example, 50 brytes indicates a level where students are likely to show considerable confidence in concrete operational tasks and, at the same time, the onset of concrete generalized reasoning. Furthermore, a *MAPS ability* of 60 brytes indicates reasonable mastery of concrete generalizations, together with an initial appreciation of logically necessary rules concerning relations in more abstract mathematical systems.

The range 25% to 75% mastery will be used in the following discussion of capabilities during each stage of mathematical development. However the overlap of stages outside this range should not be disregarded, because of the variability of both students and tasks.

The examples given in each of the following stages refer to items in various tests, mainly from the *ACER Mathematics Profile Series*. The item difficulties in brytes are shown in parentheses.

Early concrete operational stage

25% to 75% mastery range: 30 to 40 brytes

As the primary student enters the early concrete operational stage, it is reasonable to assume basic concepts such as one-to-one correspondence, and conservation of number (for details see Copeland, 1974). Basic concepts and terminology in space, time, and number (see Boehm, 1970) should have been established, together with background perceptual skills (see Frostig, 1972).

The *Operations Test* items suggest that students are developing an ability to construct and represent equivalent sets and to recognize simple patterns of relationship. Initially these are related to practical manipulations and their concrete images, e.g. $15 + 3 = \Delta + 3$ (35.0 brytes). By the end of this stage, familiar patterns of one step become generalized, but their explanations are likely to be linked to concrete situ-

ations. Inverse operations tend to be regarded as unique physical 'undoings', without much realization of their compensatory role.

Concepts of shape, space, measurement, and fractions have meaning only in concrete and specific cases at the beginning of this stage. Simple properties and relations in these areas can be explored through practical manipulations of appropriate materials.

e.g. *AM9 Length.* Item 5 (37.6 brytes)

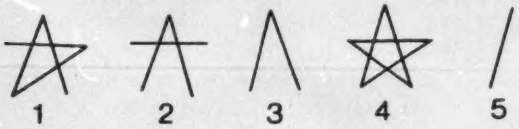
How many pieces, 20 centimetres long, can be cut from a 1 metre length of string?

A 5 pieces **C** 12 pieces
B 6 pieces **D** 8 pieces

Strategies such as one-way classification, formation of simple hierarchies, and ordering on the basis of a major attribute (e.g. age, height) should be developed by the end of this stage.

e.g. *Jenkins Intermediate Non-Verbal Test*, Part IV, Item 2 (39.5 brytes)

Which is the middle figure if the following were arranged in order?

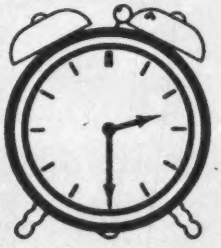


Late concrete operational stage
25% to 75% mastery range: 40 to 50 brytes

The late concrete operational stage is characterized in the *Operations Test* by items requiring students to perceive patterns of familiar concrete structure involving up to two steps. Those with symmetrical properties become generalized early in this stage and the special properties of zero and unity become sufficiently familiar to be generalized by the end of it. Compensatory manipulations such as $(987 - 321) + 321 = \Delta$ (47 brytes) are feasible in situations where the steps can be arithmetically checked, but cannot yet be considered as a general application of identity rules.

Spatial and measurement systems in one dimension become well established during this stage, but relations on two and three dimensions are still related to concrete situations. Concepts such as area and speed would be recognized informally but not yet manipulated as products and ratios of other dimensions. Simple spatial transformations, such as mirror images, become feasible towards the end of this stage.

e.g. *Space Test Unit I*, Item 28 (50 brytes)



This is the reflection of a clock in a mirror. It appears to be reading 'half past two'
The time shown by the actual clock would be

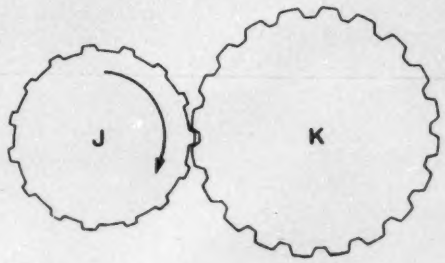
A 2.30 **C** 9.30
B 6.12 **D** 10.30

A major criterion of mastery in concrete operational thinking is student awareness of the necessary reversibility of classification and seriation strategies involving two major attributes of objects or their concrete images. Although *Jenkins Intermediate Non-Verbal Test* was developed as a norm-referenced measure of intelligence, its 80 items reflect various aspects of this criterion. It is interesting to note that the 75% mastery level on this test occurs at 50 brytes, providing an independent validation of this transition point. Student *MAPS abilities* and item difficulties for the Jenkins test are provided in the *Operations Test Teachers Handbook*, so teachers may diagnose difficulties in classification and seriation as a secondary function of this test.

The following item, which scales at 50.3 brytes, provides a further example of the logical interaction between two attributes and their inverses (clockwise/anticlockwise, faster/slower).

Measurement Test Unit II Item 3 (50.3 brytes)

J and K represent two cog wheels, with 30 and 40 teeth respectively.



If J is turning clockwise at 24 revolutions per minute, what is the direction and turning rate of K?

A clockwise at 18 revolutions per minute
B clockwise at 32 revolutions per minute
C anticlockwise at 18 revolutions per minute
D anticlockwise at 32 revolutions per minute

Early formal operational (concrete generalization) stage 25% to 75% mastery: 50 to 60 brytes

Collis's suggestion that this level of mathematical thinking should be called 'the concrete generalization' stage emphasizes that it features the ability to generalize concrete patterns but to cope with formal structures only in concrete embodiments. These aspects are clearly displayed in the *Operations Test* by the following pairs of parallel items.

Compensatory problems such as $240 + 15 = 480 + \Delta$ (53.4 brytes) are generalized to $s + t = rs + \Delta$ (54.8 brytes) without much increase in difficulty.

However the less familiar structure involved in $468 + 234 = \Delta + 468$ (55.9 brytes) is mastered only in situations where the elements can be considered arithmetically. Operations on this structure are not formally available until well into the next stage, $k + p = \Delta + k$ (66 brytes).

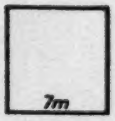
During this stage, students develop considerable understanding of spatial and measurement systems which have been developed to investigate relationships between attributes in their natural and social environment. The following items display understandings and strategies typical of the concrete generalization stage.

Space Test Unit IV, Item 22 (51.2 brytes)

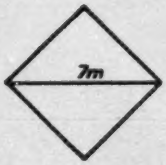
In each of these squares, one of the dimensions has been marked to represent 7 metres. Which one of the squares would have the greatest perimeter?

A

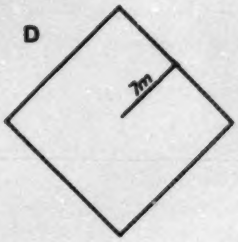
C



B

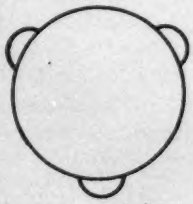


D



Space Test Unit II, Item 21 (52.5 brytes)

Peter, Sam and Sue are sitting around a table and Sam is immediately to the left of Sue. Who is sitting closest to Peter's right side?

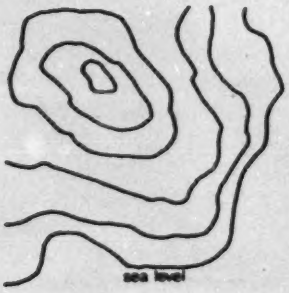


A Sam

B Sue

C It is impossible to tell without more information.

Space Test Unit IV, Item 10 (56.2 brytes)

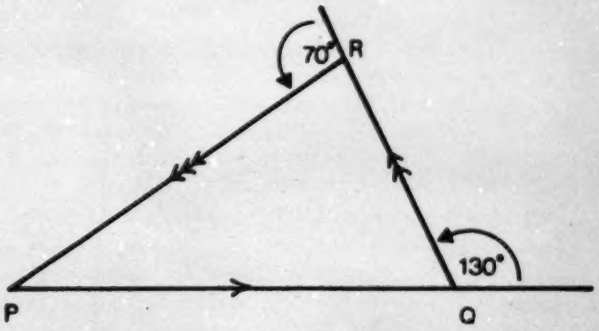


The map shows the contours of a hill near the coast. Contour lines are drawn joining points at each multiple of 10 metres above sea level.
What is the height of the hill above sea level?

A between 40 and 50 metres
B exactly 50 metres
C between 50 and 60 metres
D exactly 60 metres
E between 60 and 70 metres

Space Test Unit II, Item 32 (57.3 brytes)

This sketch gives some information about an orienteering exercise. Students set out from P, turned 130° anticlockwise at Q and 70° anticlockwise at R, returning directly to P, forming triangle PQR.



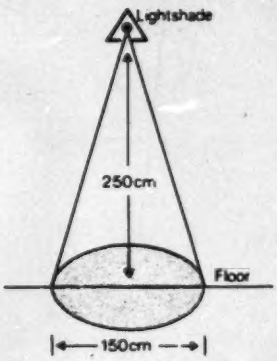
What would be the size of the angle at P between their routes out and back?

A 20°
B 30°
C 50°
D 60°
E The angle size cannot be calculated unless the route distances are known.

During this stage, students develop their ability to handle abstract ratio and rate concepts, but find it difficult to rearrange such relationships unless the manipulations are related to concrete situations. The same applies to graphical relations, especially non-linear ones, and more sophisticated spatial transformations.

Space Test Unit IV, Item 14 (57.0 brytes)

A light, with a conical shade, shines directly down onto the floor 250 cm below. Its beam illuminates a circular region of diameter 150 cm as shown in the diagram.



If the light is raised to 300 cm above the floor, the diameter of the circular region illuminated is

A 100 cm.
B 180 cm.
C 200 cm.
D 250 cm.
E 500 cm.

Late formal operational stage
25% to 75% mastery range: 60-70 brytes

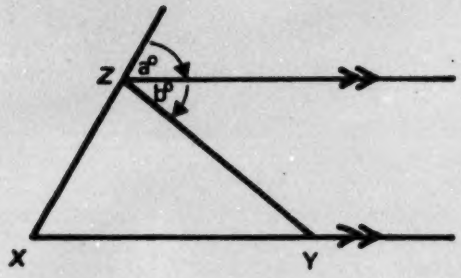
Formal operational thinking is characterized in the *Operations Test* by items with formal structure even when the elements are abstract, $f - e = \Delta - f$ (64 brytes). This also includes operations on large numbers where successive steps require systematic persistence, even if arithmetic closure is feasible, $(900 + 30) + 10 = \Delta\ (30 + 10)$ (60 brytes).

By this stage, many rules and relationships are so familiar as to be considered part of the student's reality and can be operated upon to produce all the logical transformations.

Because the student can select appropriate transformations for particular relationships, it implies a systematic view whether this is based on mathematical conventions or arises out of physical reality. With theoretical or closed systems, the formal operational student can operate on all combinations of the constituent factors or relationships.

Space Test Unit III, Item 32 (61.2 brytes)

On this diagram, the lines marked with double arrows are parallel.



The angles marked a and b will be equal

A always.

B only if ΔXYZ is right-angled.

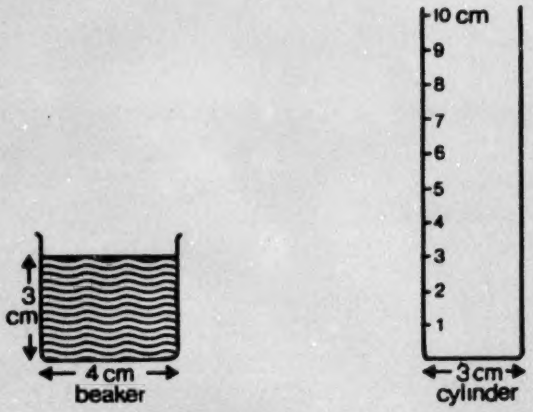
C only if ΔXYZ is isosceles, base XY.

D only if ΔXYZ is isosceles, base XZ.

In physical systems, strategies are developed sufficiently to identify relevant factors and to investigate their possible effects. Because these factors may exert combined or opposing influences, ideas of proportionality and reciprocity become very important, and their acquisition seems to be a major characteristic of students who reach this formal operational level.

Measurement Test Unit III Item 23 (62.3 brytes)

The diagram represents a cylindrical beaker, with base diameter 4 cm and a cylinder with the base diameter 3 cm. The beaker contains mercury (3 cm deep) which is poured into the cylinder.



If the markings on the cylinder are in cm, which one of these levels would be closest to the new mercury level?

A 3 cm

B 4 cm

C 5 cm

D 8 cm

Age/school year characteristics related to the MAPS measurement scale

Traditional test statistics are usually quoted in terms of one or more particular age/school year populations. The MAPS measurement scale is calibrated independently of these external characteristics but it is possible to subsequently superimpose them on the scale. This involves choosing a sample representative of the desired characteristic stratified according to the proportions within the population for each value of that characteristic. The mean *MAPS ability* of students at each value can be calculated, and thus the values of the characteristic can be assigned to those mean *MAPS ability* points on the scale.

It has not been possible to apply this procedure to superimpose accurate normative data on age or year of schooling across or within the states of Australia. Such an operation would require considerable funding and co-operative effort at a level which is not feasible at this time. However age data on students involved in Collis's research (1975, p. 168) and in the trial testing of the *Operations Test* at ACER have been combined to calculate approximate age norms. While these students were non-representative in a statistical sense, their backgrounds were quite varied, so the results provide some guidance as to expected levels of ability. Teachers wishing to make more precise normative interpretations may be prepared to use the procedures outlined above to develop local 'norms', based on mean *MAPS abilities* of students within their own school or region.

Table 8

Mean *MAPS ability* for students for ages 7 to 17 years

Age (nearest year)	Related years of schooling	Mean *MAPS ability* (nearest bryte)
7 years	2	29
8 years	3	37
9 years	4	44
10 years	5	47
11 years	6	49
12 years	7	50
13 years	8	51
14 years	9	53
15 years	10	55
16 years	11	58
17 years	12	65

The data provided in Table 8 is experimental data based on information collected by Wines (1975) and the ACER trial test results.

The standard deviations of *MAPS ability* in the ACER trial testing were of the order 6 to 7 brytes at each age level. Thus, while the mean *MAPS ability* of 12 year olds was 50 brytes and the majority (68%) ranged between 43 and 57 brytes (±1 scaled error), it is important to realize that a range encompassing 95% of 12 year olds would be very wide: 36 to 64 brytes (±2 scaled errors). While these facts emphasize the hazards of interpreting normative data, it is interesting to observe that Rasch relative mastery interpretations would suggest similar proportions of success for a group of students of mean *MAPS ability* centred on 50 brytes: ±7 brytes, 65%; ±14 brytes, 91%. It should also be emphasized that mean *MAPS abilities* for students of age 16 and 17 years relate only to those still attending senior secondary school.

Conclusions

Major emphasis has been placed on superimposing regions which reflect the operational characteristics of items calibrated on that scale, while lesser emphasis has been placed on the student characteristics of age and school year. While student maturation and yearly promotion reflect increasing ability to some extent, it has been considered more important to emphasize progress through operational stages which are more useful than age or school year in identifying the nature of appropriate learning experiences. Both types of data have been summarized in Figure 8 and this is set against the *item profiles* (pages 25-26 and the foldouts) for the *Space Test*.

The features of thinking at each operational level have been discussed in general terms and have also been exemplified by items from tests in the *ACER Mathematics Profile Series* and from associated ACER tests which have been calibrated on to the MAPS scale. The learning experiences which have been suggested are mainly process-orientated in that they can apply across various content or topic areas. These process objectives, under the general headings of BASIC PROCESSES, UNDERSTANDING PROCESSES, and APPLICATION PROCESSES, have been used in most of the *ACER Mathematics Profile Series* handbooks to group and classify items to assist in interpretation. Within these general groupings are more specific process categories such as Knowledge, Relevance, Communication, Problem Solving, Reasoning, Creativity, etc. Many teachers will be familiar with these categories which are loosely based on the cognitive processes suggested by Bloom *et al* (1970). In Table 9 they are used as the central axis of more specific process objectives which may also serve as a summary of the learning activities suggested for each operational level. With the latter purpose in mind, the process objectives are presented in two columns. The left-hand column summarizes learning experiences which should be relevant to the concrete operational student (*MAPS ability* < 50 brytes), in that they can be related to activities in the environment or with structured materials. The right-hand column summarizes learning objectives for the student who is developing an appreciation of more abstract systems in mathematics.

The learning objectives on the left-hand side also represent minimum mathematical competencies that almost every student should be able to master during schooling, to assist competent and confident participation in society. Not all students are likely to have the ability or incentive to become a formal operational thinker in mathematics, as implied by the objectives on the right-hand side. However each student should be encouraged to develop his or her reasoning power to its maximum. To this end, it is hoped that the interpretations from the *ACER Mathematics Profile Series*, using the MAPS measurement scale, will assist teachers in enhancing this development through the suggestion of appropriate learning activities.

Table 9

To ensure competence and confidence to participate in society, each pupil should develop understanding in CONCRETE situations related to the natural, physical and social environment by:	**Process objectives**	To acquire habits of effective thinking and intellectual independence, some pupils should develop understanding in ABSTRACT situations, related to structured systems which provide a framework for logical reasoning by:
	BASIC PROCESSES	
Knowing terms, units and conventions, perceiving attributes and relationships for numbers, shapes, objects, sets, positions, directions, and measures of time, money, temperature and space (1, 2, 3 dimensions).	**KNOWLEDGE**	Knowing appropriate symbols and principles, perceiving structures, properties and relationships in systems of numeration, spatial and logical relations, and complex or extended measures.
Performing computations and their inverses.	**OPERATIONS**	Manipulating symbols, relations and logical statements, their inverses and converses.
Classifying on 1 and 2 attributes; finding equivalents, opposites and complements; comparing and contrasting characteristics of numbers, shapes, objects, sets, etc.	**CLASSIFICATION**	Cross-classifying, differentiating, comparing and contrasting on two or more attributes, co-ordinating and generalizing properties and relations.
Ordering objects, numbers, measures etc. in sequences and in hierarchies; stating ordinal relationships and their inverses; finding equivalent elements in two sequences (direct proportion); making simple transformations and translations.	**ORDINATION**	Manipulating ordinal relations within systems; ordering and relating paired values and properties of variables (direct and inverse proportion); interpolating and extrapolating values and properties; translating and transforming relationships (2 or more changes).
	UNDERSTANDING PROCESSES	
Distinguishing between cardinal and ordinal uses of number in society; recognizing relevant and irrelevant attributes in grouping and ordering; recognizing symmetry and asymetry, patterns and similarities; recognizing qualitative and quantitative aspects of measurement, estimating and approximating measures.	**RELEVANCE**	Recognizing the significance of conventions in structured systems; translating from real-life to symbolic and structured systems; interpretation from structured systems back to reality, with awareness of limitations; recognizing the arbitrary nature of measurement systems, and appropriate accuracy and tolerances.
Recognizing and interpreting mathematical information of relevance in everyday life: patterns, scales, grids, plans, recipes, tables, simple graphs, networks and flowcharts.	**COMMUNICATION**	Producing and intepreting modes of representation for mathematical concepts, structures and relationships: symbolic statements, expressions and relations, co-ordinate systems, more complex graphs, networks and flowcharts.

APPLICATION PROCESSES

Applying known relationships to familiar problem situations – recognizing question and relevant data – selecting and applying appropriate operation(s) – checking answer and detecting errors.	**PROBLEM SOLVING**	Manipulating and applying literal relationships; applying thinking strategies to less familiar situations – identifying relevant data and likely relationships – considering alternative strategies – selecting and applying efficient strategy – interpreting and evaluating solution
Applying simple reasoning to interesting situations, concrete puzzles, games and codes (1-2 steps).	**REASONING**	Applying reasoning strategies to logical argument or proof; more complex puzzles, codes, etc.
Extending patterns with numbers, shapes, etc.; rearranging models, networks, maps, plans; transforming familiar conventions in practical situations and predicting effects.	**CREATIVITY (cognitive aspects)**	Extending, combining, or rearranging relationships, changing or inventing conventions of systems and predicting resulting relationships.
Recognizing relevant factors for appreciation of the advantages or disadvantages in present mathematical usage (modes of measurement, calculation, communication, problem-solving, etc.).	**ATTITUDES (cognitive aspects)**	Recognizing relevant factors for appreciation of the value and limitations of structured systems (modes of simplification of complex situations, with related assumptions about other factors, framework for logical reasoning, systematic analysis, etc.).

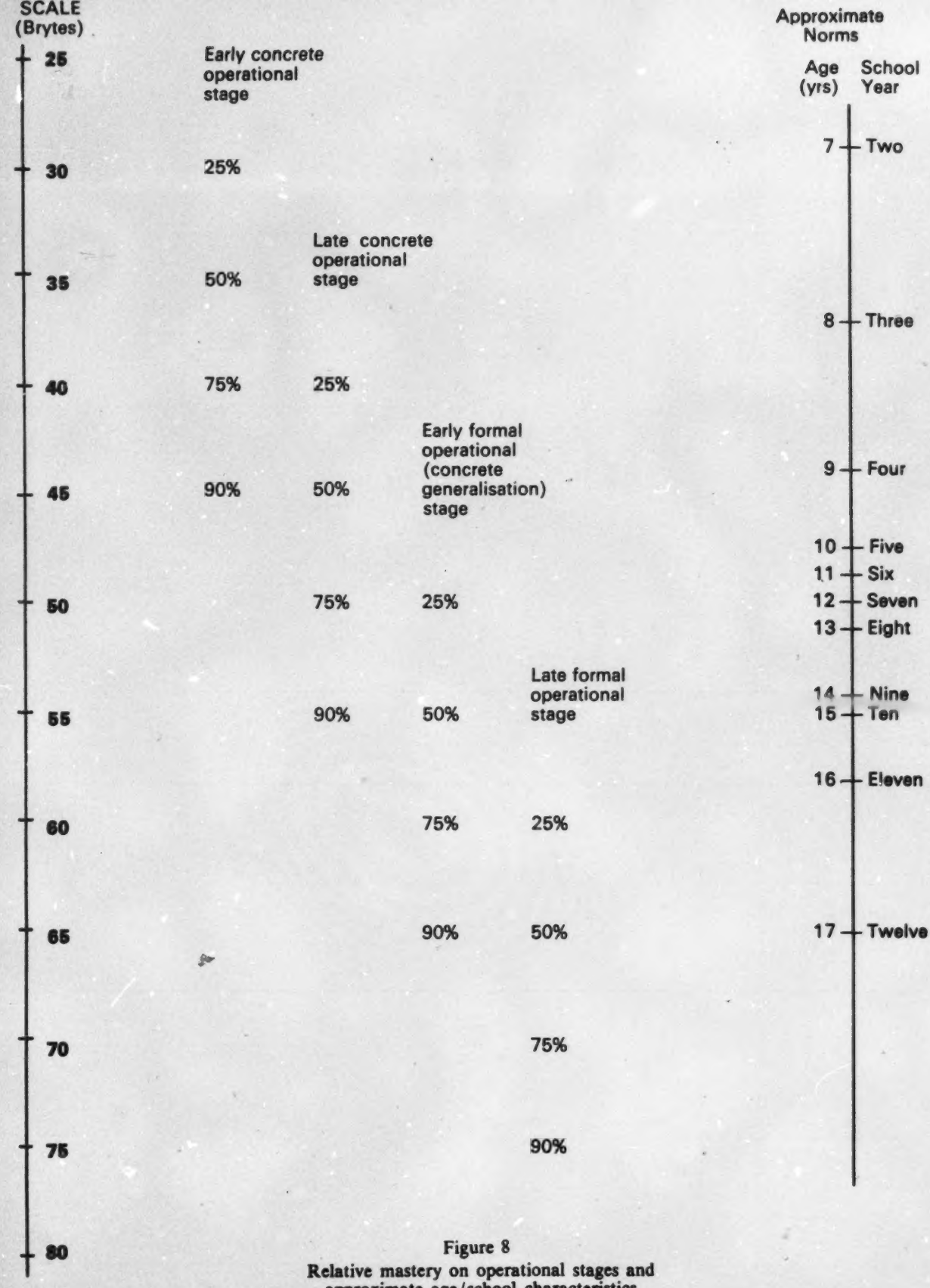


Figure 8
Relative mastery on operational stages and approximate age/school characteristics

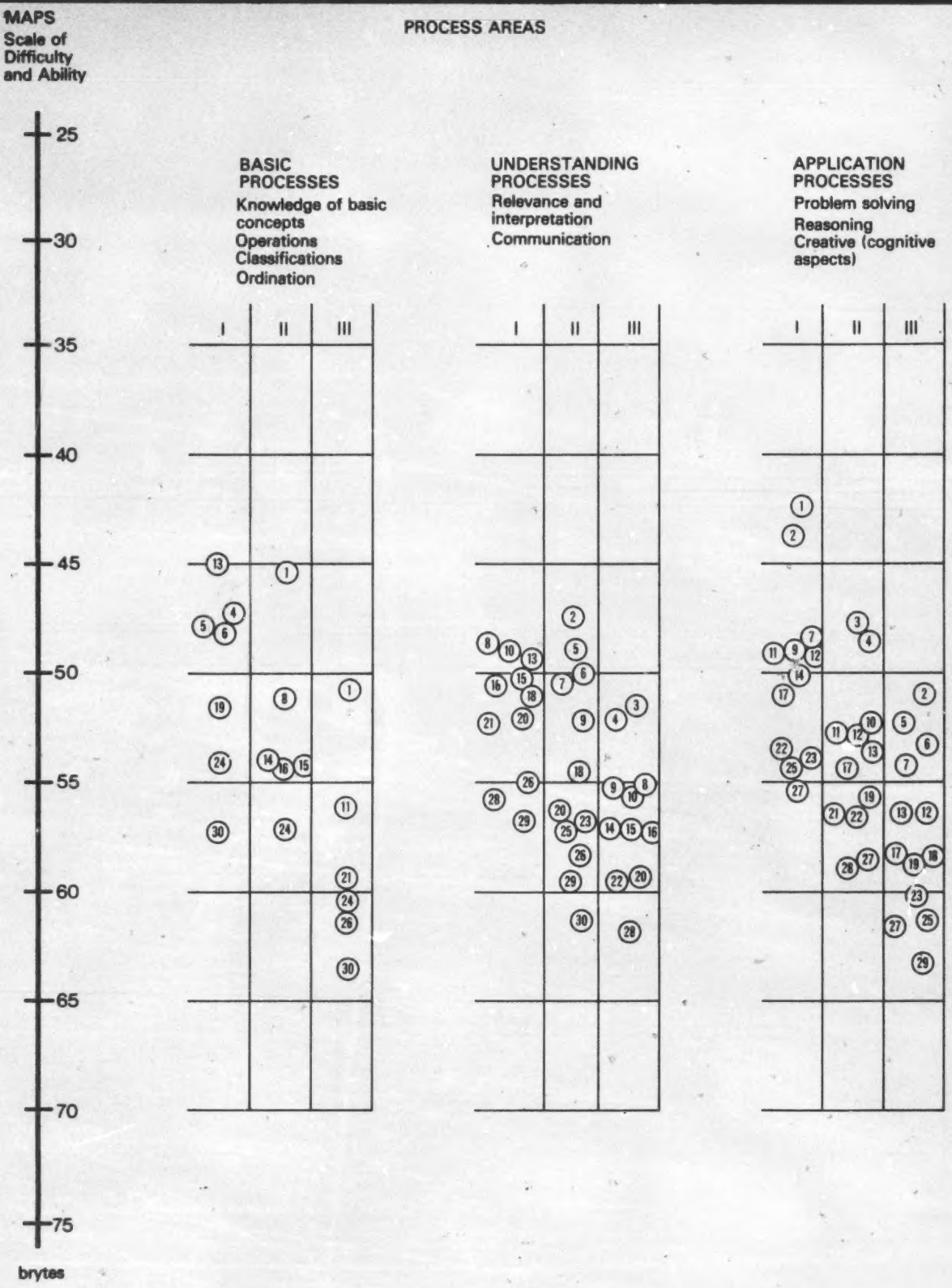


Figure 9
Number Test – item profile: Process Areas

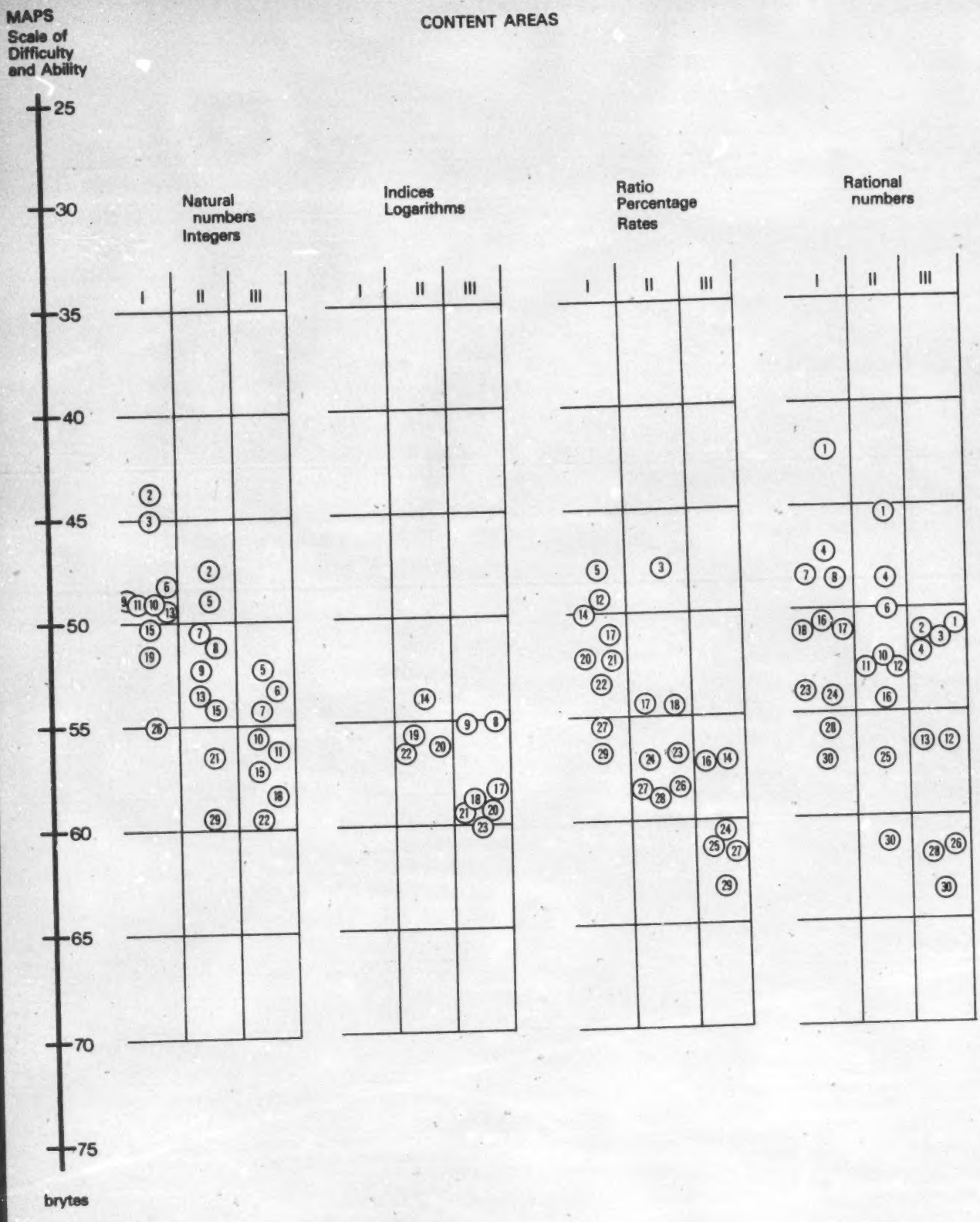


Figure 10
Number Test – item profile: Content Areas

Section IV

Results and Conversion Table

Table 10

Number Test: Item MAPS *difficulties* (*brytes*)

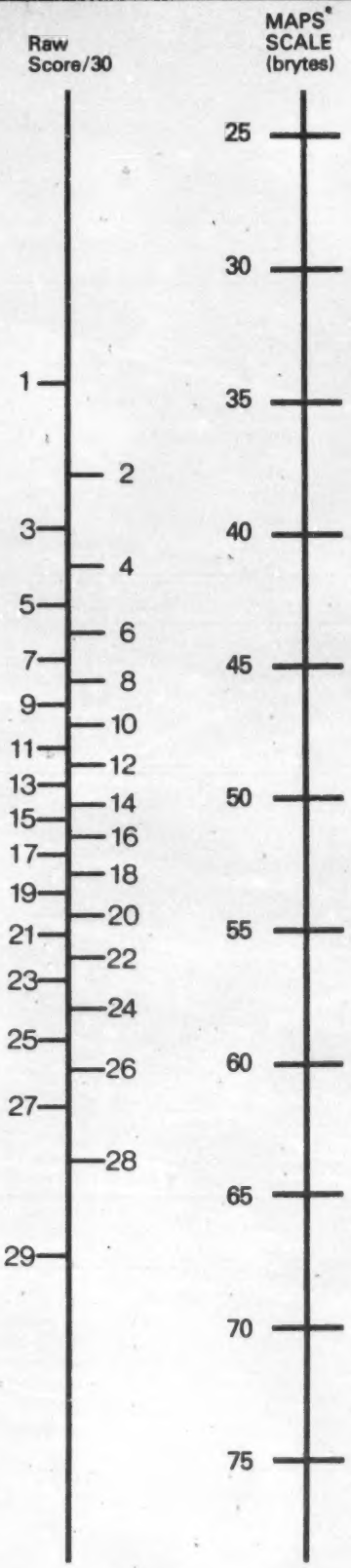
Item \ Unit	I	II	III
1	42.4	45.5	50.8
2	43.8	47.5	51.0
3	45.0	47.8	51.5
4	47.2	48.6	52.1
5	47.9	49.0	52.3
6	48.2	50.1	53.3
7	48.4	50.5	54.3
8	48.6	51.1	55.1
9	49.0	52.2	55.2
10	49.1	52.3	55.7
11	49.2	52.8	56.2
12	49.4	52.9	56.5
13	49.5	53.7	56.5
14	50.1	54.0	57.2
15	50.3	54.3	57.2
16	50.6	54.5	57.3
17	51.0	54.5	58.3
18	51.0	54.7	58.5
19	51.5	55.7	58.9
20	52.1	56.3	59.4
21	52.3	56.5	59.5
22	53.5	56.7	59.7
23	54.0	56.9	60.2
24	54.2	57.2	60.5
25	54.5	57.4	61.3
26	55.0	58.5	61.5
27	55.5	58.6	61.7
28	55.8	59.0	61.9
29	56.8	59.7	63.3
30	57.3	61.3	63.7
Mean *difficulty*	50.8	54.0	57.4
Standard deviation	3.7	4.0	3.7

CONVERSION: RAW SCORE TO MAPS *Ability*

Number Test – UNIT I (30 items)

Mean item *difficulty* = 50.8 brytes
Standard deviation = 3.7 brytes

Raw Score	MAPS *Ability* (nearest bryte)	Scaled error (brytes)
0	<34	–
1	34	4.7
2	38	3.4
3	40	2.9
4	41	2.5
5	43	2.3
6	44	2.2
7	45	2.1
8	46	2.0
9	46	1.9
10	47	1.9
11	48	1.8
12	49	1.8
13	49	1.8
14	50	1.8
15	51	1.8
16	52	1.8
17	52	1.8
18	53	1.8
19	54	1.8
20	54	1.9
21	55	1.9
22	56	2.0
23	57	2.1
24	58	2.2
25	59	2.3
26	60	2.5
27	62	2.8
28	64	3.4
29	67	4.7
30	>67	–

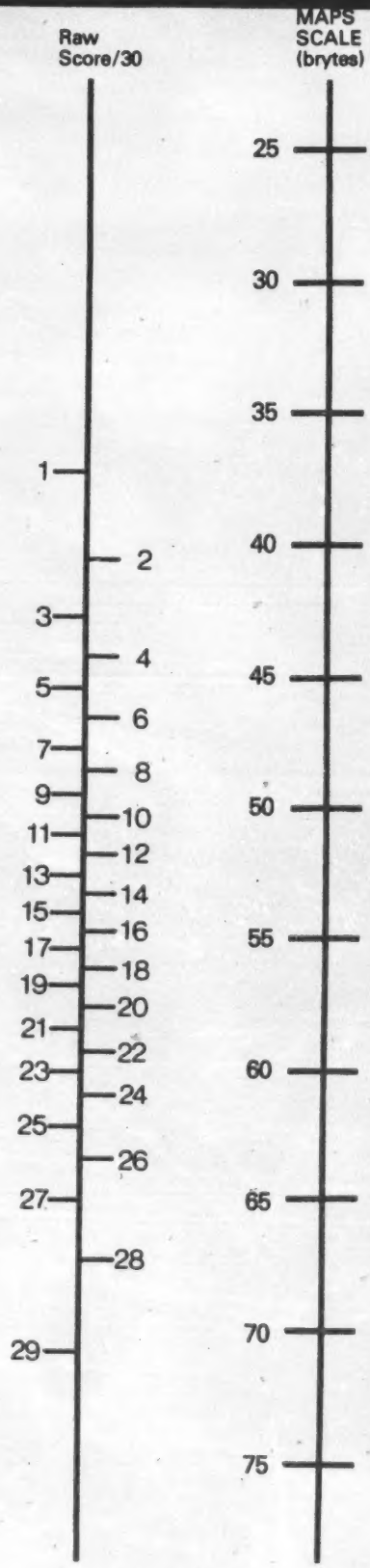


CONVERSION: RAW SCORE TO MAPS *Ability*

Number Test – UNIT II (30 items)

Mean item *difficulty* = 54.0 brytes
Standard deviation = 4.0 brytes

Raw Score	MAPS *Ability* (nearest bryte)	Scaled error (brytes)
0	<37	–
1	37	4.7
2	41	3.4
3	43	2.9
4	44	2.6
5	46	2.4
6	47	2.2
7	48	2.1
8	49	2.0
9	50	2.0
10	50	1.9
11	51	1.9
12	52	1.8
13	53	1.8
14	53	1.8
15	54	1.8
16	55	1.8
17	56	1.8
18	56	1.8
19	57	1.9
20	59	1.9
21	59	2.0
22	60	2.1
23	61	2.2
24	62	2.3
25	64	2.5
26	65	2.8
27	67	3.4
28	67	3.4
29	71	4.7
30	>71	–

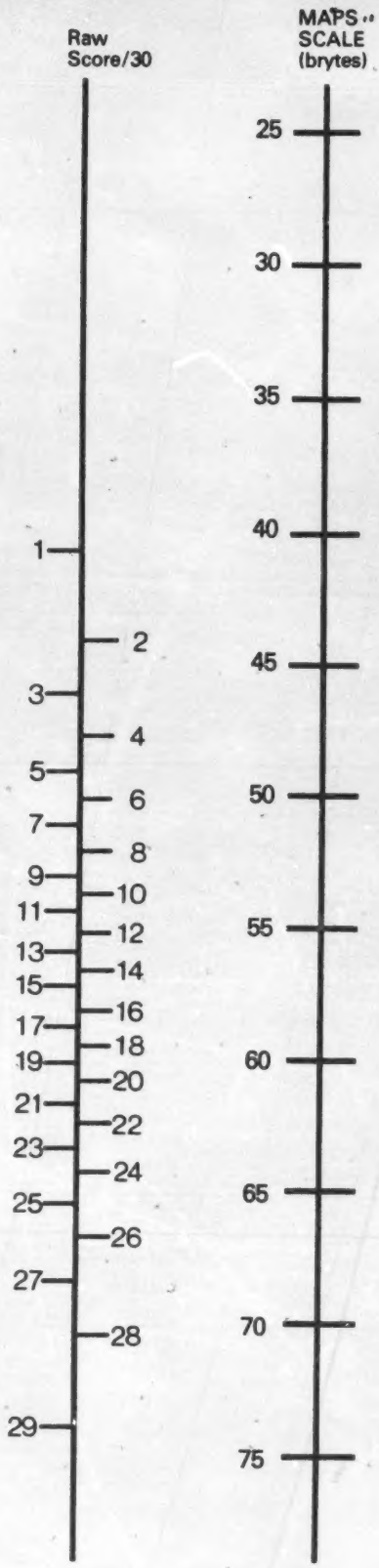


CONVERSION: RAW SCORE TO MAPS *Ability*

Number Test – UNIT III (30 items)

Mean item *difficulty* = 57.4 brytes
Standard deviation = 3.7 brytes

Raw Score	MAPS *Ability* (nearest bryte)	Scaled error (brytes)
0	<41	–
1	41	4.7
2	44	3.4
3	46	2.9
4	48	2.5
5	49	2.3
6	50	2.2
7	51	2.1
8	52	2.0
9	53	1.9
10	54	1.9
11	55	1.9
12	55	1.8
13	56	1.8
14	57	1.8
15	57	1.8
16	58	1.8
17	59	1.8
18	60	1.8
19	60	1.8
20	61	1.9
21	62	1.9
22	63	2.0
23	64	2.1
24	65	2.2
25	66	2.3
26	67	2.5
27	68	2.8
28	71	3.4
29	74	4.7
30	>74	–



Bibliography and References

ACER Mathematics Tests, (AM Series). (1971, 1974, 1976.) ACER.

Australian Science Education Project. *A Guide to ASEP.* (1974). Unit 401. (State of Victoria.)

BLOOM, B. S., HASTINGS, J. T., AND MADAUS, G. F. (1971). *Handbook of Formative and Summative Evaluation of Student Learning.* (McGraw Hill Publishing Company: New York.)

Boehm Test of Basic Concepts. Test and Teachers Manual. (1970). (The Psychological Corporation: New York.)

COLLIS, K. F. (1972). Concrete to Abstract – A new viewpoint. *The Australian Mathematics Teacher.* Vol. 28, No. 3, 113–118.

COLLIS, K. F. (1975). *A Study of Concrete and Formal Operations in School Mathematics: A Piagetian Viewpoint.* (Australian Council for Educational Research: Hawthorn, Victoria.)

CONNOLLY, A. J., NACHTMAN, W., AND PRITCHETT, E. M. (1971). *KeyMath Diagnostic Arithmetic Test Manual.* (American Guidance Service, Inc.: Circle Pines, Minn.)

COPELAND, R. W. (1974). *How Children Learn Mathematics.* (Collier-MacMillan Publishers: London.)

DOUGLAS, G. A. (1974). Test design strategies for the Rasch psychometric model. Unpublished doctoral dissertation. Department of Education, University of Chicago.

FOGELMAN, K. R. (1970). *Piagetian Tests for the Primary School.* (National Foundation for Educational Research Publishing Company: Slough, Bucks.)

FORBES, D. W. AND INGEBO, G. S. (1975). An Empirical Test of the Content Homogeneity Assumption in Rasch Item Calibration. Paper presented at the American Educational Research Association, Washington, D.C.

INHELDER, B. AND PIAGET, J. (1958). *The Growth of Logical Thinking from Childhood to Adolescence.* (Routledge and Kegan-Paul: London.)

Jenkins Intermediate Non-Verbal test, Test and Manual. (1974). ACER: Hawthorn, Vic.)

KOHLBERG, L. AND MAYER, R. (1972). Development as the Aim of Education. *Harvard Educational Review.* Vol. 42, No. 4, 449–496.

PEEL, E. A. (ed.) (1972). The Quality of Thinking in Secondary School Subjects. *Educational Review.* Vol. 24, No. 3.

PETERS, R. S. Education and Human Development. *In* DEARDEN, HIRST, AND PETERS (eds). (1972). *Education and the Development of Reason.* (Routledge and Kegan-Paul: London.)

RASCH, G. An Individualistic Approach to Item Analysis. *In* P. F. LAZARSFELD AND N. W. HENRY (eds). (1966). *Readings in Mathematical Social Sciences.* pp.89–107. (Science Research Associates Inc.: Chicago.)

SUMNER, R. (1975). Tests in Attainment in Mathematics in Schools: Monitoring feasibility study. NFER Occasional Reports, Second Series. (NFER Publishing Company: Slough, Bucks.)

WILLMOTT, A. S. AND FOWLES, D. (1974). *The objective interpretation of test performance: the Rasch model applied.* (NFER Publishing Company: Windsor, Berks.)

WINES, R. M. (1975). The Rasch approach to sample-free item analysis and person measurement – its application to item banking and diagnostic testing. Unpublished M Ed thesis.

WOODCOCK, R. (1973). Woodcock Reading Tests Manual (American Guidance Service, Inc., Circle Pines, Minn.)

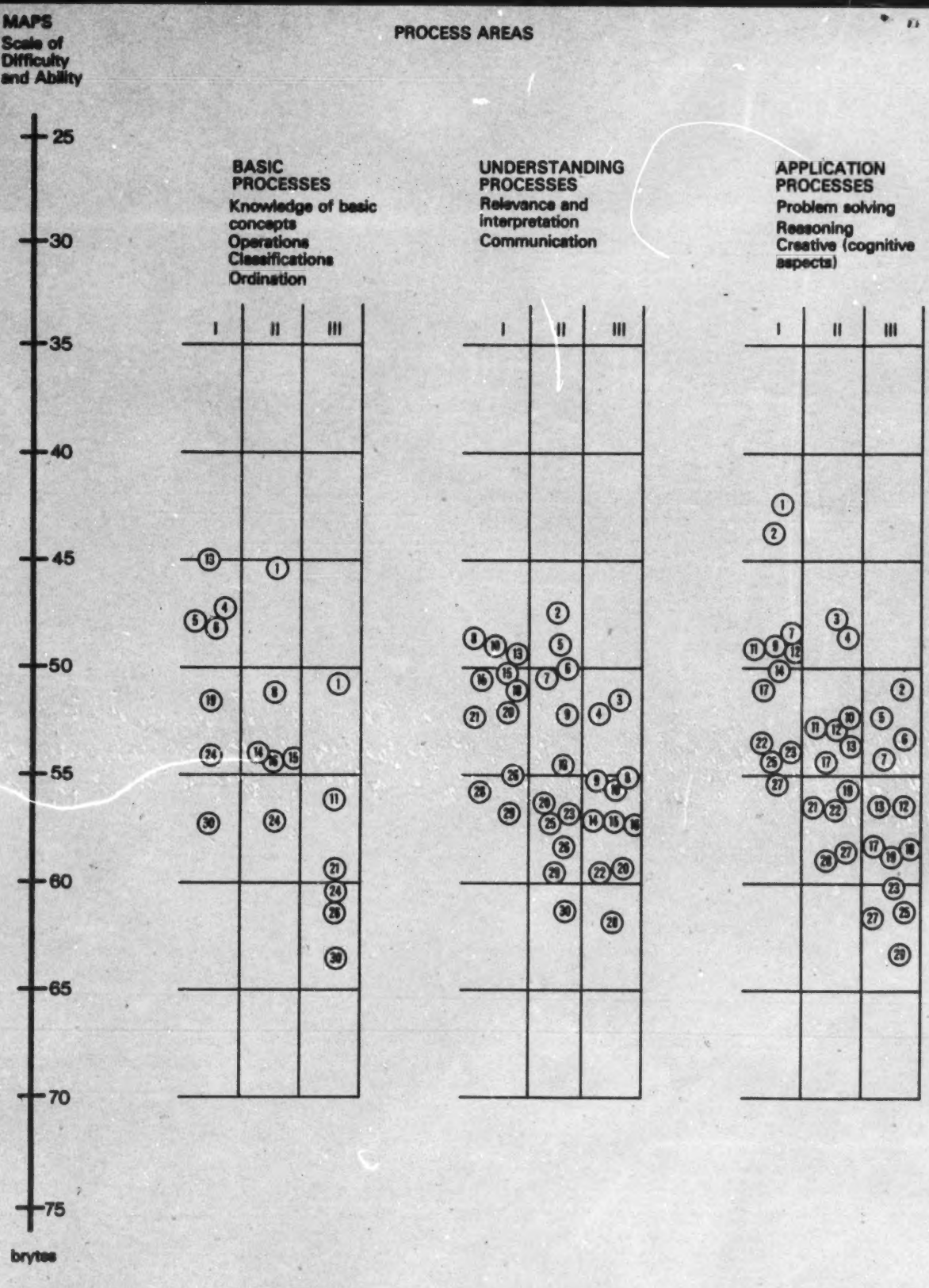
WOODCOCK, R. AND DAHL, M. (1971). A Common Scale for the Measurement of Person Ability and Test Item Difficulty. *AGS Paper No. 10.* (American Guidance Service, Inc.: Circle Pines, Minn.)

WRIGHT, B. D. (1968). Sample Free Test Calibration and Person Measurement. *Proceedings of the 1967 Invitational Conference on Testing Problems* pp.85–101. (Educational Testing Service: Princeton.)

WRIGHT, B. D. (1975). The construction and calibration of item banks and the use of calibrated items for making measurements. Report for the Board of Secondary Education, Perth, Western Australia.

WRIGHT, B. D. AND DOUGLAS, G. A. (1975). Best test design and self-tailored testing. *Research Memorandum No. 19.* Statistical Laboratory, Department of Education, University of Chicago.

WRIGHT, B. D. AND PANCHAPAKESAN, N. (1969). A Procedure for Sample-Free Item Analysis. *Educational and Psychological Measurement.* Vol. 29, No. 1, 23–48.



Number Test — item profile: Process Areas